U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

PB-280 480

The Ability of Mildly Hearing-Impaired Individuals to Discriminate Speech in Noise

Aerospace Medical Research Lab, Wright-Patterson AFB, Ohio

Prepared for

Environmental Protection Agency, Washington, D C

Jan 78

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.
EPA-550/9-78-100

2.

3. RECIPIENT'S ACCESSION NO.
PB280480

4. TITLE AND SUBTITLE
The Ability of Mildly Hearing-Impaired Individuals to Discriminate Speech in Noise

5. REPORT DATE
Jan. 1978

6. PERFORMING ORGANIZATION CODE
Project #7321, Task #03

7. AUTHOR(S)
Alice H. Suter

8. PERFORMING ORGANIZATION REPORT NO.
AMRL-TR-78-4

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Aerospace Medical Research Laboratory Wright-Patterson
Air Force Base, Ohio

10. PROGRAM ELEMENT NO.
62202F

11. CONTRACT/GRANT NO.

12. SPONSORING AGENCY NAME AND ADDRESS
EPA-ONAC
Washington D.C. 20460

13. TYPE OF REPORT AND PERIOD COVERED
Final

14. SPONSORING AGENCY CODE
EPA/ONAC

15. SUPPLEMENTARY NOTES

16. ABSTRACT

The investigation explores the relationship between hearing level at various audiometric frequencies and speech discrimination in different noise backgrounds. The study was designed specifically to test the American Academy of Ophthalmology and Otolaryngology's (AAOO) selection of a 26-decibel average of 500, 1000 and 2000 Hz as the point above which hearing handicap occurs. The AAOO method for computing hearing handicap has lately been brought into question for two primary ~~reasons: that~~ the 26-dB fence is too high, and for the exclusion of frequencies above 2000 Hz. The present study, therefore, attempted to see if there were differences among individuals whose hearing was at or better than the low fence, and if so, what factors caused or affected the differences. Forty-eight subjects were tested with two types of speech materials: the University of Maryland Test #1 which employs simple, "everyday" sentences, and the Modified Rhyme Test, a closed-set test of rhyming monosyllables. Speech stimuli were presented at 60 dBA measured at the listener's ear. The noise stimulus, a babble of twelve voices, was presented at levels of 60 to 66 dBA. Subjects were divided into three groups according to their hearing levels at 500, 1000, and 2000 Hz. One group had normal hearing at all tested audiometric frequencies and the other two had mild hearing losses in the mid-frequencies and considerable amounts of loss in the high frequencies, which is typical of noise-hearing loss. Subjects listened to both speech materials in a quiet condition and in (cont. on back)

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Hearing Impairment, Speech discrimination Noise, hearing handicap fence		

18. DISTRIBUTION STATEMENT
Limited Supply Available at EPA/ONAC
and Research Triangle Park, N.C.

19. SECURITY CLASS (This Report)
Unclassified

20. SECURITY CLASS (This page)
Unclassified

21. NO. OF PAGES
109

22. PRICE
A06 - A01

EPA Form 2220-1 (9-73)

three levels of background noise. Group scores were compared in a three-factor analysis of variance and correlations between audiometric frequencis and individual discrimination scores were performed. Conclusions are drawn about the differences between the three groups, the effect of increasing noise on the discrimination of speech by hearing-impaired individuals, the equivalence of two kinds of speech materials, and the importance of certain audiometric frequencies to the discrimination of speech. Recommendations are made for computing hearing handicap in terms of the height of the fence and the inclusion of frequencies above 2000 Hz.

★ U.S. GOVERNMENT PRINTING OFFICE: 1978—260-880/5

PB 280 480

EPA 550/9-78-100
AMRL-TR-78-4

THE ABILITY OF MILDLY HEARING-IMPAIRED INDIVIDUALS TO DISCRIMINATE SPEECH IN NOISE

JANUARY 1978

JOINT EPA/USAF STUDY

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U. S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

U.S. Environmental Protection Agency
Washington, D.C. 20460

Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base, Ohio

THE ABILITY OF
MILDLY HEARING-IMPAIRED INDIVIDUALS
TO DISCRIMINATE SPEECH IN NOISE

by
Alice H. Suter

Investigation Conducted and Report Prepared
under the Joint Sponsorship of

U.S. Environmental Protection Agency
Washington, D.C.

Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base, Ohio

January 1978

ii

SUMMARY

The purpose of the investigation was to explore the relationship between hearing level at various audiometric frequencies and speech discrimination in different noise backgrounds. The study was designed specifically to test the American Academy of Ophthalmology and Otolaryngology's (AAOO) selection of a 26-dB average of 500, 1000, and 2000 Hz, as the point above which hearing handicap occurs. The AAOO method for computing hearing handicap has lately been brought into question for two primary reasons: that the 26-dB fence is too high, and for the exclusion of frequencies above 2000 Hz. The present study, therefore, attempted to see if there were differences among individuals whose hearing was at or better than the low fence, and if so, what factors caused or affected the differences.

In designing the study, the following experimental questions were posed:

1. What is the relationship between average hearing level at 500, 1000, and 2000 Hz and speech discrimination scores in noise for individuals whose average hearing levels are at or better than the AAOO low fence?
2. Is the relationship dependent upon speech-to-noise ratio?
3. Is the relationship between average hearing level and speech discrimination scores differently described by different speech materials?
4. Which combination of audiometric frequencies best predicts speech discrimination scores?

Forty-eight subjects were tested with two types of speech materials: the University of Maryland Test #1, which employs simple, "everyday" sentences, and the Modified Rhyme Test, a closed-set test of rhyming monosyllables. Speech stimuli were presented at 60 dBA measured at the listener's ear. The noise stimulus, a babble of twelve voices, was presented at levels of 60 to 66 dBA. Subjects were divided into three groups according to their average hearing levels at 500, 1000, and 2000 Hz. Group I had normal hearing at all frequencies, Group II had mean hearing levels of 13.4 dB in the mid-frequencies, and Group III, mean hearing levels of 24.7 dB in the mid-frequencies. Both groups II and III had considerable amounts of loss in the high frequencies, which is typical of noise-induced hearing loss. Subjects listened to both speech materials in a quiet condition and in three levels of background noise. Group scores were compared in a three-factor analysis of variance, and correlations between audiometric frequencies and individual discrimination scores were performed.

Results of the investigation provided answers to the above questions as follows:

1. Significant differences were found in mean speech discrimination scores among all of the three groups, showing that within the "normal" area under the 26-dB fence, there was considerable variation in the ability to discriminate speech in noise.

2. The relationship between average hearing level and speech discrimination scores proved to be dependent upon speech-to-noise ratio. The discrimination scores of all three groups decreased as the speech-to-noise ratio became lower, and the differences between groups increased.

3. Mean scores were similar for the two kinds of materials in the two intermediate noise conditions, but not in quiet, or in the most difficult noise condition. However, the two materials were equally effective at delineating differences among the three groups in the various conditions.

4. Correlational tests revealed that frequency combinations that included frequencies above 2000 Hz were significantly better predictors of speech discrimination scores than the combination of 500, 1000, and 2000 Hz.

PREFACE

According to a Public Health Survey* in 1962, 8.4% of the U.S. population, or nowadays approximately 18 million people have hearing levels of 25 dB or greater at the average audiometric frequencies of 500, 1000 and 2000 Hz. This hearing level is at the point of beginning hearing handicap, as it is defined by the medical profession today. Many more individuals have hearing levels that are less severe, but nevertheless complain of difficulties in understanding speech, especially in a background of noise. Many of these 18 million or more individuals have suffered their hearing losses as a result of exposure to noise.

On the authority of the Noise Control Act of 1972 the Environmental Protection Agency is charged with conducting research on the effects of noise. The resulting information is to be used for developing and refining criteria, which in turn is used for setting standards and regulations, advising other Federal agencies, giving technical assistance to local communities, and educating the general public, all for the general purpose of protecting the public against the adverse effects of noise.

The research described in this report was undertaken to assess the functional abilities of individuals with hearing levels of approximately 25 dB in the mid-frequencies, as well as those with hearing levels less severe. By doing so, the Agency would obtain information on the masking effects of certain levels of environmental noise on a substantial portion of the population. More specifically, the research was designed to test the adequacy of the 25-dB mid-frequency demarcation point as the beginning of hearing handicap. The results would, therefore, be of interest to those who develop rules or guidelines for medico-legal purposes that are based upon a definition of hearing handicap. They would also be of interest to physicians, audiologists, public health specialists, and Federal agency personnel who have traditionally used the 25-dB, mid-frequency hearing level to differentiate between "normal" and "impaired" hearing.

The research was carried out in the Biodynamics and Bioengineering Division of the Aerospace Medical Research Laboratory at Wright-Patterson Air Force Base, Ohio, in support of Project 7231, Work Unit 03.** It was conducted by Ms. Alice Suter of the Environmental Protection Agency's Office of Noise Abatement, under the supervision of Dr. H. E. von Gierke of the Air Force, under the auspices of an Inter-agency Agreement between the two agencies. In addition to Dr. von Gierke the following individuals served as reviewers of this report:

Dr. William Burns	Dr. Karl Kryter	Mr. Karl Pearsons
Laleham-on-Thames	Stanford Research Institute	Bolt Beranek and Newman, Inc.
England	Menlo Park, California	Canoga Park, California

---

* U.S. Department of Health, Education and Welfare, Public Health Service, Hearing levels of adults by age and sex, United States 1960-1962.

** The voluntary informed consent of the subjects used in this research was obtained as required by AFR 80-33.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1
  Statement of the Problem 3

II. HISTORY AND SURVEY OF THE LITERATURE 5
  Historical Efforts to Assess Hearing Handicap 5
  Support for the AAOO Rule 8
  Criticism of the AAOO Rule 9
  Other Methods of Assessing Hearing Handicap 10
    Veterans Administration Method 10
    Social Adequacy Index 11
    Self Evaluation 11
  Experimental Attempts to Characterize Hearing Impairment or Handicap 12
    Tests in Quiet 12
    Tests in Noise 12
    Other Distortions 14
  Rationale for the Present Study 15

III. PROCEDURES 17
  Subject Selection 17
  Speech Materials 23
    UM Test #1 23
    MRT 24
  Masking Noise 25
  Relationship of Speech and Noise Stimuli 25
    Speech Level 25
    Speech-to-Noise Ratio 26
  Sound Field 27
  Listening Mode 27
  Instrumentation 28
    Equipment 28
    Measurement 28
    Calibration 30
    Hearing Protective Devices 30
  Method of Presentation 31
  Analysis of the Data 32
    Corrections for Guessing 33
    Three-Factor Design 33

IV. RESULTS 35
  Analysis of Variance 35
  Correlational Tests 39
  Further Ad Hoc Comparisons 43

V. DISCUSSION 46
  Statistical Analysis 46
    Analysis of Variance 46

TABLE OF CONTENTS (Cont)

Chapter Page

Other Tests on the Data 47
Comparisons of Results with Other Studies 48
Differences Between Groups 48
Effect of Speech-to-Noise Ratio 49
Similarity of Speech Materials 50
Comparison with Other Correlational Studies 50
Monaurality 53
The Speech Frequencies 53
Fences 54

VI. SUMMARY AND CONCLUSIONS 58
Summary 58
Conclusions 60
Recommendations 60

Appendix A Pilot Experiments 62
Instrumentation 62
Pilot Study #1 62
Pilot Study #2 62
Pilot Study #3 64

Appendix B Consent of Volunteer 68

Appendix C Instructions to Subjects (MRT) 69

Appendix D Speech Materials 70
Modified Rhyme Test 70
University of Maryland Test #1 71
Lists B through J

Appendix E Raw Data
Subjects' Hearing Levels and Discrimination Scores 80

References 83

LIST OF TABLES

Table | Page

1. Means, standard deviations, and ranges of better-ear hearing levels in dB for each audiometric frequency. 20

2. Means, standard deviations, and ranges of better-ear hearing levels in dB for various combinations of frequencies. 21

3. Age and educational levels of the three groups. 22

4. Attenuation values in dB for hearing protectors as reported by U.S. Air Force (1973). 31

5. Mean discrimination scores in percent correct for MRT before and after corrections for guessing, and for UM Test #1. 34

6. Mean scores and standard deviations in percent correct responses. 36

7. Summary of analysis of variance to determine effects of hearing loss, speech-to-noise ratio, and speech materials on speech discrimination scores. 37

8. Tests on significant interactions of hearing loss with speech-to-noise ratio, and speech materials with speech-to-noise ratio. 38

9. Correlational matrix among six combinations of audiometric frequencies and discrimination scores in various speech-to-noise ratios. 42

10. Matrix displaying significance of differences among correlations of various frequency combinations and speech discrimiation scores averaged over the four experimental conditions. 44

A-1. Pilot Study #1. Mean scores and standard deviations in percent correct of ten normal-hearing subjects for various single and combined lists of UM Test #1. 63

A-2. Pilot Study #2. Mean scores and standard deviations in percent correct on the Modified Rhyme Test of thirteen normal-hearing subjects in monaural and binaural listening modes. 65

A-3. Pilot Study #3. Individual scores on MRT and UM Test #1 in a variety of speech-to-noise ratios. 66

LIST OF FIGURES

Figure		Page
1	Mean better-ear hearing levels and ranges of the three experimental groups.	19
2	Block diagram of equipment.	29
3	Mean percent correct responses of the three groups as a function of speech-to-noise ratio. Scores are averaged across the two speech materials.	40
4	Mean percent correct response in the two speech materials as a function of speech-to-noise ratio. Scores are averaged across the three groups.	41
5	Mean speech discrimination scores in percent correct as a function of speech-to-noise ratio. Groups II and III have been combined and partitioned according to whether subjects have more or less hearing loss than the median (47 dB) at the average of 2000, 3000, and 4000 Hz. Group Y = < 47 dB, Group Z = > 47 dB.	45

CHAPTER I

INTRODUCTION

In today's complex society, hearing and understanding speech is by far the most important function of the hearing mechanism. Warning signals and the sounds of nature are also important to the human listener, depending upon individual circumstance and preference. But it is generally agreed that when loss of hearing becomes a handicap, it is because an individual can no longer adequately hear or understand speech.

Hearing loss from exposure to noise has been recognized through the ages, especially in conjunction with metalworking and gunpowder (Burns, 1973). It became a pervasive occupational condition with the advent of the industrial revolution. Today, there are approximately 14 million American workers in the production industries, 70% of whom are exposed to noise levels of 85 dBA and above (BBN, 1974). Most of these individuals will incur some amount of hearing loss, however small in some cases, if they remain in their noisy jobs over a working lifetime (EPA, 1974).

Before any significant attempts were made to prevent occupational hearing loss, the condition was recognized as job-related and therefore compensable under workmen's compensation laws. In the early part of the century only "acoustic trauma" was compensated, but in the 1950's noise induced hearing loss of gradual origin was recognized as a compensable occupational disease (Newby, 1964; Ginnold, 1974). Since compensation preceded prevention, it is not very surprising that compensation formulas found their way into damage-risk criteria, and thus the concepts of compensation and prevention became confused (Suter and von Gierke, 1975; von Gierke, 1975). The risk of hearing loss from noise was stated in terms of the percentage of workers expected to incur compensable losses when exposed to certain levels of noise for certain amounts of time (Rudmose, 1957).

The terms disability, handicap, and impairment were used almost interchangeably until 1965, at which time the American Academy of Ophthalmology and Otolaryngology (AAOO) decided to make a long-needed distinction (Davis, 1965).

1. Disability: actual or presumed inability to remain employed at full wages.
2. Impairment: a deviation or change for the worse in either structure or function, usually outside of the range of normal.

3. Handicap: the disadvantage imposed by an impairment sufficient to affect one's personal efficiency in the activities of daily living.

That these terms are still confused is evident in the 1974 proposed standard for noise exposure of the U.S. Department of Labor's Occupational Safety and Health Administration (OSHA). The proposed OSHA standard uses the same criterion for impairment that the AAOO has used for handicap, (OSHA, 1975).

In the ensuing controversy over the proposed OSHA standard new questions about occupational hearing loss have been raised. It has become clear that there are legal and administrative aspects of the issue on the one hand, and scientific aspects on the other. Examples of the former are the differentiation between prevention and compensation, and the decisions as to the amount of hearing loss that should be prevented, and the amount of compensation that should be awarded. Examples of the latter are the estimation of hearing loss resulting from noise exposure, and the decision as to how much hearing loss is a handicap. This last point should be decided primarily on the basis of hearing loss for speech.

Currently, the most widely used method for dealing with hearing loss, both for compensation and preventive purposes, is the AAOO method (Lierle, 1959). This method designates lower and upper cutoff points or "fences" at 26 dB and 93 dB (converted to the ANSI 1969 standard for audiometric zero), respectively, for the averaged frequencies 500, 1000, and 2000 Hz with 1-1/2% handicap assigned to each decibel of hearing level between those points.

The AAOO rule and its rationale have been criticized recently on a variety of grounds: (a) that the rule is inappropriate for preventive criteria, (b) that the "low fence" is too high, (c) that the rule discounts the value of high-frequency hearing, (d) that it fails to take the noisy aspect of day-to-day living into account, and (e) that it is not based on sufficient scientific evidence (Kryter, 1963, 1973, and 1975; Niemeyer, 1967; Kuzniarz, 1973; EPA, 1974; von Gierke, 1975; Meyer, 1975; Thomas, 1975; UK-Member Body, 1975). The reason for the persisting confusion between preventive and compensation criteria is not entirely clear, except, as it was pointed out above, that the early emphasis on industrial hearing conservation was to prevent compensable hearing loss. Some other nations (such as the United Kingdom) have made the distinction, even though the AAOO's definition of beginning handicap has found its way into the International Organization for Standardization's "Assessment of occupational noise exposure for hearing conservation purposes" (ISO, 1971). This confusion between prevention and compensation has been brought to the attention of OSHA by the Environmental Protection Agency (EPA, 1975), and by Meyer (1975), Thomas (1975), and von Gierke (1975) in formal testimony during OSHA-sponsored public hearings.

The height of the 26 dB-fence has been criticized by Kryter (1973). Based on calculations using the Articulation Index, Kryter maintained that an individual with an average hearing level of 26 dB (at 500, 1000, and 2000 Hz) can correctly repeat only 90% of sentences at a normal conversational level and 50% of monosyllabic words at a "weak" conversational level. The calculations assume a quiet background, and a distance of one meter between talker and listener. Kryter recommended a low fence of 15 dB for the frequencies 500, 1000, and 2000 Hz.

With respect to the frequencies used in predicting hearing handicap, the National Institute of Occupational Safety and Health (NIOSH, 1972) has dropped 500 Hz and substituted 3000 Hz in its criteria for the assessment of occupational hearing impairment, thereby affirming the importance of high frequency hearing. The decision was based largely on the research of Harris, Haines, and Meyers (1960), Kryter, Williams, and Green (1962), and Niemeyer (1967). (While no distinction is mentioned, the NIOSH criteria pertain to preventive rather than compensation criteria.)

Kryter, Williams, and Green (1962), Niemeyer (1967), and Kuzniarz (1973) among others, have demonstrated the importance of high frequency hearing in conditions of background noise. Since some amount of background noise is very common in everyday listening conditions, these investigators have proposed that frequencies higher than 2000 Hz be included in the averaging process for the prediction of speech discrimination scores.

Statements issued at the time of the AAOO's deliberations (DeForest and Lierle, 1955; and Lierle 1959) defined hearing for speech as "the ability to identify spoken words or sentences under average everyday conditions of normal living . . ." but did not specify the noise conditions under which such sentence or word identifications could occur, or the degree of correct identification that was considered acceptable. Although one assumes that the selection of a formula must have been guided by some kind of scientific evidence, such evidence is not apparent. According to the AAOO's Subcommittee on Noise (Lierle, 1959): "These principles are based on current medical opinion."

Statement of the Problem

The present study was undertaken with the intent of examining the AAOO low fence, both in terms of the 26-dB cutoff level and the use of the simple, unweighted average of 500, 1000, and 2000 Hz. The object was to see whether individuals whose hearing impairments are at this criterion or better are indeed "not handicapped." In other words, does the AAOO method implicitly classify some individuals as normal who yet have considerable difficulty in understanding speech?

In order to investigate this issue subjects were selected whose hearing in the mid-frequencies (500, 1000, and 2000 Hz) fell in the vicinity of or better than the AAOO low fence. Any amount of loss above 2000 Hz was permitted in the experimental subjects, since these losses are not considered in the AAOO formula. The effect of these losses was central to the study. Speech discrimination scores were obtained for sentence and monosyllabic material in several different levels of background noise in order to assess the influence of increasingly difficult listening conditions.

A dramatic cutoff in hearing level between little and great difficulty in speech discrimination could not be expected, just as the 26-dB fence cannot be viewed as a magical turning point. Loss of hearing is a complex phenomenon, with great possibilities for subtle physiological and behavioral differences. Statements of pure-tone acuity as measured by standard clinical techniques cannot fully describe it. However, pure-tone thresholds are a simple, objective method for measuring hearing level, and they continue to be used for medico-legal purposes despite certain shortcomings. Therefore, in this experiment pure-tone thresholds have been related to speech discrimination scores under a variety of conditions, and some attempt has been made to describe the functional abilities of individuals according to their pure-tone thresholds.

In summary, this study investigated the adequacy with which the AAOO rule predicts the point of beginning handicap through (a) application of the fence at 26 dB, and (b) exclusion of frequencies above 2000 Hz. To this end the following specific questions were posed:

1. What is the relationship between average hearing level at 500, 1000, and 2000 Hz and speech discrimination scores in noise for individuals whose average hearing levels are at or better than the AAOO low fence?
2. Is the relationship dependent upon speech-to-noise ratio?
3. Is the relationship between average hearing level and speech discrimination scores differently described by different speech materials?
4. Which combination of audiometric frequencies best predicts speech discrimination scores?

HISTORY AND SURVEY OF THE LITERATURE

Historical Efforts to Assess Hearing Handicap

The early experimental work on the intelligibility of speech was not done to assess hearing handicap, but rather to study and improve communication systems. A vast body of knowledge was developed by such well-known investigators as Davis, Fletcher, French, Steinberg, and Stevens. This information was utilized in the early attempts to assess hearing handicap and still is used today. These early investigators spoke of "hearing loss for speech" since the distinctions between disability, impairment, and handicap had not yet been made. Fletcher (1929) developed the first well-known method for assessing hearing loss for speech, based on loudness and intelligibility data. The entire audible range from 0 dB to 120 dB (re ASA 1951 audiometric zero) for the averaged frequencies 500, 1000, and 2000 Hz was divided into degrees of loss with a slope of 0.83% loss per dB. The simple (unweighted) average of 500, 1000, and 2000 Hz was chosen because it correlated well with results of unamplified live voice and phonograph record tests using digits. The slope was later modified to yield an even 0.8% per dB and hence the time-honored "Fletcher Point-Eight Rule." According to Davis, the rule was not meant for purposes of compensation because there was no threshold of handicap and the ceiling was virtually unattainable (Davis, 1970).

In 1939, the American Medical Association's (AMA) House of Delegates of the Council on Physical Therapy approved a recommendation for standardizing tests and preparing a method for estimating percentage loss of hearing. The result was a "Tentative Standard Procedure for Evaluating the Percentage of Useful Hearing Loss in Medicolegal Cases" (Carter, 1942). The proposed standard ascribed percentage values of hearing loss to octave intervals between the frequencies 256 and 4096 Hz, and to 10-dB intervals from 10 dB to approximately 95 dB (re ASA 1951 audiometric zero). One plotted an audiogram, connected the hearing levels at the various frequencies, selected the percentage values immediately above the connected lines and summed them. Binaural hearing loss was computed by weighting the better ear seven times the value of the poorer ear. The standard was endorsed by the AMA in 1942.

This first AMA method was prepared by Bunch, Fowler, and Sabine (as reported by Fowler, 1947) and was derived almost entirely from a method developed by Sabine (1942). Sabine's method was in turn based upon studies conducted by Knudsen (1923) and Fletcher (1929), that described the number of discriminable units for both pitch and loudness within the speech range. It was also based on the work of Steinberg and Gardner (1940) who measured the abilities of both normal-hearing and hearing-impaired persons to understand speech at various suprathreshold levels. In addition it was influenced to some extent by a method developed by Fowler (1942) which relied on Fowler's own clinical experience as an otologist, and on a combination of the experimental data on frequency filtering by Fletcher, Steinberg, and others at Bell Laboratories (Fletcher, 1929).

The 1942 method was revised in 1947 (Carter, 1947) and later became known as the Fowler-Sabine or AMA method. The new method more closely resembled Fowler's 1942 formula in that percentage values were ascribed to discrete frequencies rather than to octave intervals, and the frequency weighting was uniform at all hearing levels, namely 15%, 30%, 40%, and 15% for 500, 1000, 2000, and 4000 Hz, respectively. In the old AMA method, frequency weighting varied according to intensity.

Although the new formula was somewhat simpler than the 1942 version, it was still not very popular, and many physicians preferred the old Fletcher Point-Eight Rule. According to Davis (1973), "it was essential to simplify the more accurate but complicated Fowler-Sabine scale in order to gain acceptance. Otologists just wouldn't use the complicated table...." In addition to complaints about the complexity of the AMA method, it was felt that the weightings ascribed to the different frequencies were somewhat arbitrary (Davis, 1971). Another objection was that although the formula worked fairly well with conductive losses, it was not appropriate for individuals with sensori-neural losses that were predominantly in the high frequencies (De Forest and Lierle, 1955). Consequently, another method was proposed in 1959 (Lierle, 1959), which was adopted by the AMA in 1961. It was prepared by the Subcommittee on Noise of the American Academy of Ophthalmology and Otolaryngology, and it became known as the AAOO method or rule.

The AAOO method used the simple average of 500, 1000, and 2000 Hz with a "low fence" or normal cutoff at 15 dB (ASA; or 26 dB re ANSI audiometric zero) and a "high fence" or total loss cutoff at 82 dB (ASA; 93 dB ANSI), with 1-1/2% impairment (or later handicap) for each dB between the two fences. The better ear was given five times the weight of the poorer ear. This method was purportedly based on the ability of the hearing-impaired individuals to hear speech. In the words of the Subcommittee (Lierle, 1959):

> Ideally, hearing impairment should be evaluated in terms of ability to hear everyday speech under everyday conditions... The ability to hear sentences and repeat them correctly in a quiet environment is taken as satisfactory evidence of correct hearing for everyday speech. Because of present limitations of speech audiometry, the hearing level for speech should be estimated from measurements with a pure tone audiometer. For this purpose, the Subcommittee recommends the simple average of 500, 1000 and 2000 cps... If the average hearing level at 500, 1000 and 2000 cps is 15 dB [ASA] or less, usually no impairment exists in the ability to hear everyday speech under everyday conditions.

According to Davis (1973), the Subcommittee determined, on the basis of clinical evidence, that an average hearing level at 500, 1000, and 2000

Hz of approximately 16 dB (ASA) was the point at which individuals begin to have difficulty hearing sentences in quiet and seek medical help for their hearing problems.

Glorig and his colleagues (see Glorig and Baughn, 1973) conducted a self-assessment poll in conjunction with the Wisconsin State Fair hearing survey of 1954. Responses to the question "Is your hearing good, fair, or poor?" were correlated with median hearing levels. Individuals who reported that their hearing was "good" had median hearing levels of 10 to 16 dB (ASA; averaged over 500, 1000, and 2000 Hz). However, Glorig and Baughn point out that the responses were clearly age-dependent, with younger people rating themselves by more stringent criteria than the older people.

The Subcommittee's report (Lierle, 1959) made no mention of experimental data in support of its decision to change the Fowler-Sabine method, although there were studies conducted in the decade between the AMA and AAOO rules that favored the simple average of 500, 1000, and 2000 Hz as being the most important indicator of "hearing for speech." Carhart (1946) found that the AMA method and the simple average of the three frequencies 512, 1024, and 2048 Hz correlated equally well with speech reception thresholds for bi-syllabic words in a variety of hearing-impaired cases. Although the AMA method was found to be a slightly better predictor of speech reception for subjects with marked high tone losses, there was less variability associated with the 3-frequency method. For this reason and for practical considerations, Carhart favored the 3-frequency method.

As mentioned above, otologists had continued to use Fletcher's Point-Eight rule, mainly because of its simplicity. As a result of additional studies, Fletcher first reaffirmed the simple 3-frequency average, and then proposed a more complex method employing weighted frequencies from 250 to 8000 Hz. The former method (Fletcher, 1950), based entirely on loudness calculations, was validated against three sets of speech reception data. These data consisted of the averaged speech reception thresholds (level of 50% correct responses) for digits, spondees, phonetically balanced (PB) words, and sentences, transmitted by a variety of means (phonograph records, calling directly through the air, and voice attenuated by an audiometer). Fletcher found that while the average of 500, 1000, and 2000 Hz was a good predictor of speech reception for relatively flat losses, the average of the best two of these three frequencies was more often correct, especially for sloping losses, and hence the widely used "Fletcher Average" or "Two-Frequency Average" that still is in clinical use today.

Fletcher's next method of assessing hearing loss for speech (Fletcher and Galt, 1950; and Fletcher, 1952) was based on calculations of the Articulation Index. This formula employed the following weights: 250 Hz = 0.04, 500 Hz = 0.13, 1000 Hz = 0.23, 2000 Hz = 0.30, 4000 Hz = 0.25, 8000 Hz = 0.05. The study's purpose was to match transmission systems to hearing-impaired ears. It does not seem to have influenced current thinking on the assessment of hearing handicap as much as his 1950 formula.

Two other investigations are likely to have given impetus to the AAOO method. Harris, Haines, and Myers (1956) compared six different pure-tone methods for predicting "hearing loss for speech" as measured by thresholds of intelligibility (50% correct) for PB words. The authors concluded that the best predictor was a combination of a multiple regression method (including 500 to 6000 Hz) and the 3-frequency (500, 1000, and 2000 Hz) average. After an error in calculation was pointed out, the authors concluded that the simple average of 500, 1000, and 2000 Hz was the most satisfactory method (Harris et al, Erratum, 1956).

In a later study, Quiggle, Glorig, Delk, and Summerfield (1957) showed that the average of 500, 1000, and 1500 Hz was the best predictor of hearing loss for speech as measured by spondee word thresholds. The authors simplified the method by substituting 2000 for 1500 Hz (since 1500 Hz is rarely tested), but cautioned that spondee words were not "fully representative of the speech sounds a man must 'understand' to communicate verbally."

Although the psychoacoustical and statistical techniques used in these attempts at characterizing hearing handicap may have been fairly sophisticated, the actual assessment of hearing for speech was incomplete, and often primitive by today's clinical standards. Such techniques as calling words and digits through the air give little quantifiable information on an individual's capacity to discriminate speech sounds. In addition, threshold measures (a 50% criterion), for PB's as used by Harris et al (1956), or spondees as used by Carhart (1946) and Quiggle et al (1957), give little information on one's ability to understand speech in a variety of everyday conditions. All of the above experiments were presumably carried out in quiet since none mentions a noise background.

Support for the AAOO Rule

Some investigators have continued to support the AAOO rule since its acceptance by the AMA in 1961. Davis (1971) stated, "The AAOO rule now enjoys considerable legal prestige by its incorporation into many rules or even State laws relating to compensation for hearing handicap, from whatever the cause." It has been enthusiastically supported by Glorig (see Discussion of Part I and Summing-up in Robinson, 1971; Glorig and Baughn, 1973). Glorig and Baughn (1973) cited three studies in defense of the use of 500, 1000, and 2000 Hz. The first, by Ward, Fleer, and Glorig (1962), showed relatively small differences for speech discrimination in quiet between subjects with sloping hearing losses above 2000 Hz and those with normal hearing. In the second, Myers and Angermeier (1972) tested speech discrimination in subjects with a variety of hearing losses. The authors found large amounts of scatter when the audiometric average of 500, 1000, and 2000 Hz, and discrete frequencies of 2000 Hz and 3000 Hz were related to speech discrimination scores in quiet and noise, and they concluded that no pure-tone audiometric index could explain the variance among their individual listeners. The third study, by Murry and Lacroix (1972) investigated speech discrimination in noise of individuals with hearing losses above 2000 Hz. The authors found that none of the pure tone predictors tested (average of 500, 1000, and 2000 Hz; 1000, 2000, and 3000 Hz; or 3000 Hz alone) correlated well with speech discrimination scores, but that the average of 1000, 2000, and 3000 Hz correlated slightly better than the other two. They also found that individuals with hearing losses above 2000 Hz scored about 5 to 10 percentage points more poorly than normal-hearing subjects on the discrimination test used (the Modified Rhyme Hearing Test). In short, these studies lend no support to the 500, 1000, and 2000 Hz formula, contrary to the implication of Glorig and Baughn.

Criticism of the AAOO Rule

At present there is considerable criticism of the AAOO rule in the scientific community, especially as it applies to preventive criteria. The Department of Labor held public hearings on the proposed standard for occupational exposure to noise in 1975, which provided a forum for this kind of criticism. The most frequent complaint was that the simple average of 500, 1000, and 2000 Hz penalizes persons with noise-induced hearing losses by giving equal weight to these three frequencies, and by ignoring frequencies above 2000 Hz, even though noise-exposed individuals sustain most of their loss in the higher frequencies (von Gierke, 1975; Kryter, 1975; Thomas, 1975). In the criteria for an occupational noise exposure standard developed by NIOSH (1972) the 500, 1000, and 2000 Hz average is rejected in favor of 1000, 2000, and 3000 Hz because of the importance of higher frequencies for understanding speech in everyday conditions. Similarly, the British Standard for Assessment of Occupational Noise Exposure (BS 5330) has recently changed to an average of 1000, 2000, and 3000 Hz, and these same frequencies are incorporated into the U.K.-Member Body's recommendation for revision of the ISO 1999 (1975). The importance of the higher frequencies for understanding speech has been frequently reported in the literature.

Kryter (1963 and 1973) has maintained that the 26-dB (ANSI) fence is too high, especially in a background of noise. From calculations based on the Articulation Index, he predicted that an individual with a 26-dB hearing loss at 500, 1000, and 2000 Hz can correctly repeat only 90% of sentences at a normal conversational level (55 dB long-term rms), and 50% of monosyllabic words at a weak conversational level (50 dB long-term rms). The calculations assume a quiet background and a distance of one meter between talker and listener. Kryter advocated a 15-dB fence if the frequencies 500, 1000, and 2000 Hz are used. Even Davis (1971), who has been a strong supporter, admitted that the AAOO rule may be harsh at the low fence. He has also stated that it is not appropriate for "steep" audiograms (Davis, 1970), although he has continued to defend the rule over the years.

The AAOO rule has also been criticized on the grounds that the criterion for understanding speech, namely hearing sentences and repeating them correctly in a quiet environment, is not representative of real life. Kryter (1973) summarized this point as follows:

> It is not obvious why the noise-deafened ear should not be considered impaired nor the individual handicapped when losing the ability the normal person has to understand: (a) individual words, the unexpected message, the unfamiliar name, or the important telephone number; (b) the weaker-than-normal intensity speech that can occur because the talker drops his voice level, or the distance between the talker and listener is greater than one meter or so and the talker is using a normal conversational level of effort for the quiet; or (c) speech in the presence of everyday noise, at a party or conference when several people are talking, etc.

Although undistorted speech in quiet is used in many experimental and in most clinical evaluations, Harris (1965) has pointed out that this condition is characteristic of not more than half of our everyday listening conditions. Niemeyer (1967) and Kuzniarz (1973) have also stressed the universality of background noise and its detrimental effects on individuals with high frequency hearing loss.

Other Methods of Assessing Hearing Handicap

Veterans Administration Method

Although pure-tone audiometry has been the primary method used for medico-legal and damage-risk purposes, speech audiometry has been used for some time in certain systems. The Veterans Administration (VA) has rated hearing impairment for many years either by a combination of speech reception threshold (SRT) for spondee words and speech discrimination scores for PB words, or by pure-tone thresholds. Normal limits of hearing are defined by an SRT of less than 26 dB (ANSI), a discrimination score of higher than 92%, and pure-tone thresholds of better than 40 dB (ANSI) for all audiometric frequencies from 250 - 4000 Hz and better than 25 dB (ANSI) for at least four frequencies, (VA, 1976). Hearing losses that exceed these amounts are not necessarily eligible for compensation, even though they are no longer considered normal. Compensation is usually awarded on the basis of speech audiometry, but in those cases where only pure tone thresholds are used it can be awarded for any of the following conditions (VA, 1976): hearing loss in both ears of 50 dB or greater in any of the frequencies 500, 1000, or 2000 Hz; average hearing level in both ears of 38 dB or greater at 500, 1000, and 2000 Hz; hearing loss in one ear of 75 dB or greater in any of the frequencies 500, 1000, or 2000 Hz; or average hearing level in one ear of 500, 1000, and 2000 Hz of 58 dB or greater (all values re ANSI zero).

Social Adequacy Index

The Social Adequacy Index (SAI) was developed as a tool for predicting improvement by a hearing aid or by the fenestration operation, and for assessing social adequacy for medicolegal purposes. The SAI was based on the speech discrimination scores, using Harvard PB words, of normal and hearing-impaired listeners that would be predicted at three different listening levels: faint speech at 55 dB (sound pressure level), average speech at 70 dB, and loud speech at 85 dB. The SAI represented the average of discrimination scores over the three conditions.

Davis (1948) developed a chart where the SAI could be estimated from just two audiometric tests, the speech reception threshold in dB, and the discrimination score at about 35 dB above SRT. Difficulty in social situations was thought to begin at an SAI of 67, when about two-thirds of the PB monosyllables would be understood for an average of the three conditions. This point occurred between an SRT of 28 dB (converted to ANSI zero for speech audiometry) if the discrimination score was perfect, and a discrimination score of 70% if the SRT was 2 dB (ANSI) or better. The threshold of social adequacy, the point at which one could barely "get by" was an SAI of 33. This point was judged to occur between a speech reception threshold of 46 dB (converted to ANSI zero for speech audiometry) if the discrimination score was perfect, and a discrimination score of 35% if the SRT was 2 dB or better. Although the scheme was a logical one, Davis (1970) offers two possible reasons why it did not work very well in practice. One was the fact that speech discrimination tests were not considered as accurate as tests for speech reception threshold. The other was an admitted lack of knowledge about the relationship between the understanding of connected speech, and its component parts (sound frequencies, phonemes, and syllables).

Self Evaluation

There have been many attempts to assess hearing handicap through questionnaires or self-evaluation techniques. Notable examples are studies by Nett, Doerfler, and Matthews (1959), High, Fairbanks, and Glorig (1964), and Atherley and Noble (1971). The investigators expressed reservations with pure-tone audiometry as an adequate predictor of handicap since two people with an identical pure-tone impairment will often suffer different degrees of handicap (High et al, 1964; Atherley and Noble, 1971). These methods, however, have failed to produce quantitative results. Too many variables besides hearing loss are involved. The answers appear to be dependent upon age (Glorig and Baughn, 1973; Merluzzi and Hinchcliffe, 1973), on occupation (Simonton and Hedgecock, 1953), and on a number of other factors as brought out by Nett et al (1959). Although the study by Nett et al was an ambitious attempt to assess handicap by self-evaluation, the authors failed to recommend a usable verbal model based on the resulting data, and the study was never published.

Experimental Attempts to Characterize Hearing Impairment or Handicap

Tests in Quiet

The relationship between audiometric frequency and the understanding of speech in hearing-impaired listeners has been studied for over thirty years. Most of the earlier tests were conducted in quiet (Utley, 1944; Fletcher, 1950; Quiggle, Glorig, Delk, and Summerfield, 1957; Mullins and Bangs, 1957; and Ward, Fleer, and Glorig, 1962). The work of Fletcher, Quiggle et al, and Ward et al has been discussed above.

Utley (1944) compared the hearing threshold levels of severely hard-of-hearing children, and their losses as computed by five different pure-tone methods of assessment, to discrimination scores for vowels and consonants. She found that all five pure-tone methods correlated well with discrimination scores, as did hearing levels at 512, 1024, and 2048 Hz. However, the study is not applicable to mildly hearing-impaired individuals since many of the children had no residual hearing at 2048 Hz, and even fewer had hearing above that point. Consequently, the author did not attempt to correlate thresholds above 2048 Hz with discrimination scores.

Mullins and Bangs (1957) compared pure-tone thresholds of hearing-impaired veterans to speech discrimination scores on Harvard PB word lists. Of the pure-tone thresholds 250, 500, 1000, 2000, 3000, and 4000 Hz, the investigators found that the frequencies 2000 and 3000 Hz were the best predictors of speech discrimination scores. They also found that the steepness of the slope between 500, 1000, and 2000 Hz, which they called the "index of inferred masking" for these three frequencies, yielded the highest correlation with discrimination ability of any of their measures, although none of the correlations was high enough to achieve statistical significance.

Quist-Hanssen and Steen (1960) tested three pure-tone methods of assessment against scores for monosyllables, disyllables, digits, and "context" speech (in Norwegian), in subjects with noise-induced hearing loss. The authors found that all three methods overestimated speech reception scores. Although the averaged frequencies 500, 1000, and 2000 Hz came closer to predicting speech scores than the other two methods, the authors concluded that speech audiometry should be performed in these cases.

Tests in Noise

During the 1960's and 1970's investigators began to examine the effects of background noise, presumably because noise is characteristic of many everyday listening conditions. Most of these investigations showed that good hearing at and above 2000 Hz was necessary to overcome the adverse effects of masking noise (Kryter, Williams, and Green, 1962; Ross, Huntington, Newby, and Dixon, 1965; Acton, 1970; Elkins, 1971; Lindeman, 1971; Aniansson, 1973; Kuzniarz, 1973; and Dickman, 1974).

One of the first, and largest of these studies was that of Kryter, Williams, and Green (1962), which correlated audiometric thresholds with speech discrimination scores for the Harvard sentences and monosyllables. The authors grouped a large population of sensori-neural hearing loss subjects according to the frequency at which the hearing loss began (for example 1000 Hz and above, or 2000 Hz and above, etc.). They found that the frequencies 2000, 3000, and 4000 Hz were the best indicators of speech discrimination in their experimental conditions and concluded that formulas that do not take into account frequencies above 2000 Hz are not appropriate predictors for the understanding of speech in "realistic acoustic environments."

Ross, Huntington, Newby, and Dixon (1965) also studied speech discrimination in a background of noise in an attempt to relate discrimination scores to a combination of an "exogenous distortion" (noise) and a variety of "endogenous distortions" (abnormal difference limen functions, reduced linear range, and audiometric configuration). The authors wished to test the hypothesis that suprathreshold distortions contribute to speech discrimination problems over and above the amount that would be explainable by frequency filtering. Although the multiple distortion hypothesis was not validated, the study did indicate the value of the higher audiometric frequencies (2000 and 4000 Hz), for speech discrimination in quiet. But these two frequencies did not predict the relative discrimination shift (the differences between scores in quiet and scores in noise) as well as 500 Hz.

A study of speech discrimination by Elkins (1971) showed that hearing-impaired subjects performed predictably more poorly than normal-hearing subjects, both in quiet and in noise, except for the most difficult noise condition, where the mean difference between groups was much smaller than expected. Like Ross et al, Elkins found that the frequencies 2000, 3000, and 4000 Hz were good predictors of speech discrimination in quiet, but significant correlations between audiometric threshold and speech discrimination in moderate levels of noise were lacking. Both Elkins and Ross et al presented the speech material at a level of 40 dB above each subject's SRT, with the masking stimulus at a fixed level in relation to the speech. This technique may help to explain the smaller than expected differences between normal and hearing-impaired groups in noise. It may also help to explain the lack of clear cut results in the studies by Myers and Angermeyer (1972) and Murry and Lacroix (1972).

Presenting the stimuli at a level of 40 dB above SRT is usually done in an attempt to find the asymptote of a subject's performance-intensity (P-I) curve, or "PB max." While this procedure is useful for evaluating potential performance under amplification, it is not the most appropriate method of describing the performance of hearing-impaired individuals in everyday, unamplified conditions. In everyday conditions, people are forced to listen at various levels along the P-I curve, not just at the point of maximum performance. If the investigators had presented the stimuli at a fixed level for all subjects, greater differences between normal and hearing-impaired groups would most likely have occurred. For example, a presentation level of 40 dB above audiometric zero would most likely produce a 100% correct response from a normal listener. A presentation level of 40 dB above an SRT of 24 dB in a hypothetical case of hearing impairment, (which would be 64 dB above audiometric zero) would also be likely to produce a discrimination score of 100%. If, however, the speech were presented at the same level to both listeners (+40 in one case and +16 in the other) the discrimination scores might be quite different.

Studies by Acton (1970), Lindeman (1971), Aniansson (1973), Kuzniarz (1973), and Dickman (1974) presented both the speech and noise at fixed levels regardless of the individual subjects' hearing acuity. Each of these investigators used subjects with predominantly high-frequency hearing losses who would be considered either normal or only mildly impaired when assessed by the 500, 1000, 2000 Hz average, with the possible exception of Lindeman's group, whose audiograms were not reported. Acton studied discrimination of monosyllables in a background of pink noise that had been filtered so as to roll off 6 dB per octave in the higher frequencies. He found that subjects whose losses included 2000 Hz were decidedly more impaired than those whose losses began above that point. Lindeman compared audiometric thresholds of persons with noise-induced hearing loss to their ability to identify monosyllables (Dutch) in backgrounds of "cocktail party" noise. He found that in a speech-to-noise ratio of +10 (speech 10 dB above noise), the most important audiometric frequency was 2000 Hz. Aniansson evaluated Swedish PB monosyllables in a background of traffic noise plus, in some conditions, competing radio voice and live voice. He concluded that the frequencies 3000 and 4000 Hz are just as important as 500 and 1000 Hz when estimating speech discrimination in an "everyday milieu." Kuzniarz studied the effect of white and predominantly low frequency noise on Polish monosyllable and sentence discrimination. Dickman used the Central Institute for the Deaf (CID) "everyday" sentences in three different noise backgrounds. Both of these latter investigators found significant differences in scores between normal-hearing and hearing-impaired individuals. They also concluded that hearing acuity above 2000 Hz is critical for speech discrimination in noisy conditions.

Other Distortions

Since noise is not the only distortion to which everyday speech is subject, some investigators chose to examine other distortions, such as poor articulation, speeded speech, and filtering. A study of sentence intelligibility for speeded speech led Harris, Haines, and Myers (1960) to reconsider the value of frequencies higher than 2000 Hz in hearing-impaired listeners. The authors concluded that near-normal hearing at 3000 Hz is essential for sentence intelligibility if the speech is speeded. Harris (1965) hypothesized that distortions which may have little adverse effect when experienced singly can produce a much more serious effect when combined. This theory prompted him to examine a number of speech distortions including atypical accents, interruptions, reverberation, and three degrees of speeding. Correlations with pure-tone audiograms caused him to recommend an average of 1000, 2000, and 3000 Hz as the best predictor of everyday speech, which he estimated is distorted by sources other than noise about 50% of the time. Some of the previously mentioned experimental work also employed some forms of speech distortion as well as noise. Kryter et al (1962) introduced low-pass filtering (attenuating the speech above 2000 Hz), Acton (1970) and Aniansson (1973) introduced mild reverberation, and as mentioned above, Aniansson's study included competing speech in some conditions.

Rationale for the Present Study

Although the majority of experimental evidence points to the importance of high frequency hearing for understanding speech, further research was needed specifically to test certain conditions. As Davis (1973) suggested, the AAOO assumptions needed to be validated in an experiment utilizing "everyday" speech in an "everyday" noise background. Subjects were needed whose hearing levels were in the vicinity of or better than the 26-dB low fence.

Most of the researchers mentioned earlier in this chapter did not present the speech material in "everyday" environments. The earlier experiments were conducted in quiet (Carhart, 1946; Harris et al, 1956; Mullins and Bangs, 1957; Quiggle et al, 1957). Later, most of the studies that employed noise presented the stimuli under earphones rather than in the sound field (Kryter et al, 1962; Ross et al, 1965; Elkins, 1971; Lindeman, 1971; Myers and Angermeyer, 1972; Murry and Lacroix, 1972; and Dickman, 1974). Conditions were somewhat more lifelike in the experiments of Harris (1965) and Aniansson (1973) where the stimuli were recorded in the sound field and then presented to the subjects through earphones. However, Harris' reverberation time of five seconds was considerably greater than that of most rooms, which according to Nabelek and Pickett (1974a) is more likely to be one second or less. Of all the speech discrimination studies of mildly-impaired persons discussed above, only Acton (1970) and Kuzniarz (1973) presented the speech and noise material in a live sound field, with reverberation times of 0.5 and 0.2 second, respectively. Even those experiments could have been somewhat closer to "everyday" conditions if sentences in addition to or instead of monosyllables had been used.

The need for lifelike experimental conditions in assessing hearing handicap was stressed by Webster, Davis, and Ward (1965) in their comments on Harris' experiment (1965). The authors suggested that the distortions that Harris had chosen were not "everyday" speech distortions: electronic chopping with a 50% duty cycle and 8 interruptions per second, reverberation time of 5 seconds, and speedups of 250 to 345 words per minute. Moreover, they criticized the study for not using noise as a listening condition, and for presenting the stimuli at 40 dB above speech reception threshold. Instead, they recommended that hearing handicap be assessed under the following conditions:

1. "everyday patients" with high-frequency hearing losses

2. "everyday noise," low-frequency masking

3. "everyday talkers" such as those used by Harris
4. "everyday distortions," 1 to 2 seconds reverberation time, and a background of meaningful babble
5. "everyday listening levels," not 40 dB above each patient's threshold.

The present experiment was designed to satisfy the above conditions as closely as possible. The experimental conditions will be discussed in detail in the next chapter. Under these everyday environmental conditions, the relationship between hearing levels and speech discrimination scores was explored. Also investigated was the dependency of this relationship on increasing levels of background noise, the ability of two different speech materials (sentences and rhyme words) to describe the above relationships, and the ability of various audiometric frequency combinations to predict speech discrimination scores.

CHAPTER III

PROCEDURES

This chapter will address the selection of subjects and the setting of experimental conditions. Those conditions include speech materials, masking noise, the relationship between the speech and noise stimuli, sound field, and mode of listening (monaural as opposed to binaural). Instrumention will also be discussed, as well as the method of stimulus presentation, and the selected methods of data analysis.

Two types of speech materials, a closed set test of monosyllabic words--the Modified Rhyme Test (MRT)--and an open set test of sentences--the University of Maryland Test #1 (UM Test #1)--were presented to forty-eight subjects. The subjects were divided into three equal groups according to their better-ear average hearing level for the frequencies 500, 1000, and 2000 Hz. Mean average hearing levels were as follows: Group I 2.5 dB, Group II 13.4 dB, and Group III 24.7 dB. The tests were conducted in a mildly reverberant room (T = 0.625 sec.), in four different noise conditions with the speech stimuli delivered at a fixed level of 60 dBA measured at the subject's ear. The speech-to-noise ratios were slightly different for the two different speech materials. The speech-to-noise ratios for the MRT were quiet (Q), 0, -3, and -6 dB; and for the UM Test #1 were Q, -1, -3, and -5 dB, the minus designation indicating that the noise level was higher than the speech level. Experimental conditions were determined on the basis of three pilot studies which are discussed in Appendix A.

Subject Selection

Subjects were recruited through the clinical and screening facilities at Wright-Patterson Air Force Base and the Springfield Air National Guard in Ohio, by advertisements in local newspapers, and by word-of-mouth. Both normal-hearing and hearing-impaired subjects were recruited. Testing was accomplished in one visit, and subjects were paid $15. Those with hearing losses were tested by bone conduction so that those with conductive losses could be eliminated. Also, those with losses too severe or too mild to meet the hearing loss criteria were eliminated. Subjects were screened informally in order to eliminate those with problems of articulation or dialect, which might interfere with scoring procedures on the sentence discrimination task where the subject responded orally. Air-conduction and bone-conduction testing was performed with a Grason-Stadler 1701 audiometer with TDH 49 earphones, in a room that met the ANSI (1960) specifications for audiometric test rooms.

Subjects who qualified were classified into one of three groups according to the following criteria:

Group I, the normal or control group, could have better-ear hearing levels in the averaged mid-frequencies 500, 1000, and 2000 Hz no worse than 8 dB (ANSI, 1969) with no frequency from 250-6000 Hz worse than 20 dB.

Group II could have better-ear average hearing levels in the mid-frequencies from 10 to 18 dB.

Group III could have better-ear average hearing levels in the mid-frequencies from 20 to 28 dB.

No restrictions were placed on presence, absence, or amount of high-frequency hearing loss in groups II and III, but since most of the subjects had noise-induced hearing losses, there were considerable amounts of loss in the 3000 to 6000 Hz range.

Figure 1 shows mean better-ear hearing levels and ranges for the frequencies 250 through 6000 Hz for the three groups. Table I shows standard deviations and ranges of hearing level for each audiometric frequency for the three groups. Table II shows the average hearing levels, standard deviations, and ranges for a number of possible frequency combinations. As a result of the study's design, inter-subject variability was fairly low in the low and middle frequencies, but considerable variability was allowed in the higher frequencies.

Table III shows means, standard deviations, and ranges for the factors of age and educational level. A total of 48 subjects were selected, 16 in each group. No attempt was made to control for sex since it has been shown that there is no difference in speech perception between male and female listeners (Silverstein, Bilger, Hanley, and Steer, 1953).

Subjects were required to be within the ages of 18 and 56, and to have at least an 8th grade education so as to have acquired basic language skills. The upper age limit was imposed in order to minimize the contribution of presbycusis. Investigators have found a decline in speech discrimination as a function of aging, despite controlling for actual hearing level, (a summary of these studies is found in CHABA, 1977). The effect appears to be greater with more difficult speech material (Goetzinger, Proud, Dirks and Embrey, 1961). Feldman and Reger (1967) found that discrimination scores for phonetically balanced (PB) words decreased by approximately 5% per decade after age 50 in a population that was not controlled for hearing acuity, (which decreased about 10 dB per decade in the mid-frequencies and about 15 dB per decade at 4000 Hz). (PB words are more difficult than the materials selected for the present study). Blumenfeld, Bergman, and Millner (1969), using the Fairbanks Rhyme Test, found that correlations of scores with age were much higher with subjects over age 60 than for those below that age. Bergman (1971) tested mildly hearing-impaired subjects in quiet, with the "CHABA" sentences (the same material that was selected for the present study), and found that there was little degradation in scores until age 80.

The effect of aging also appears to be greater when the speech is presented in a background of noise than in quiet (Jerger, 1973). Mayer

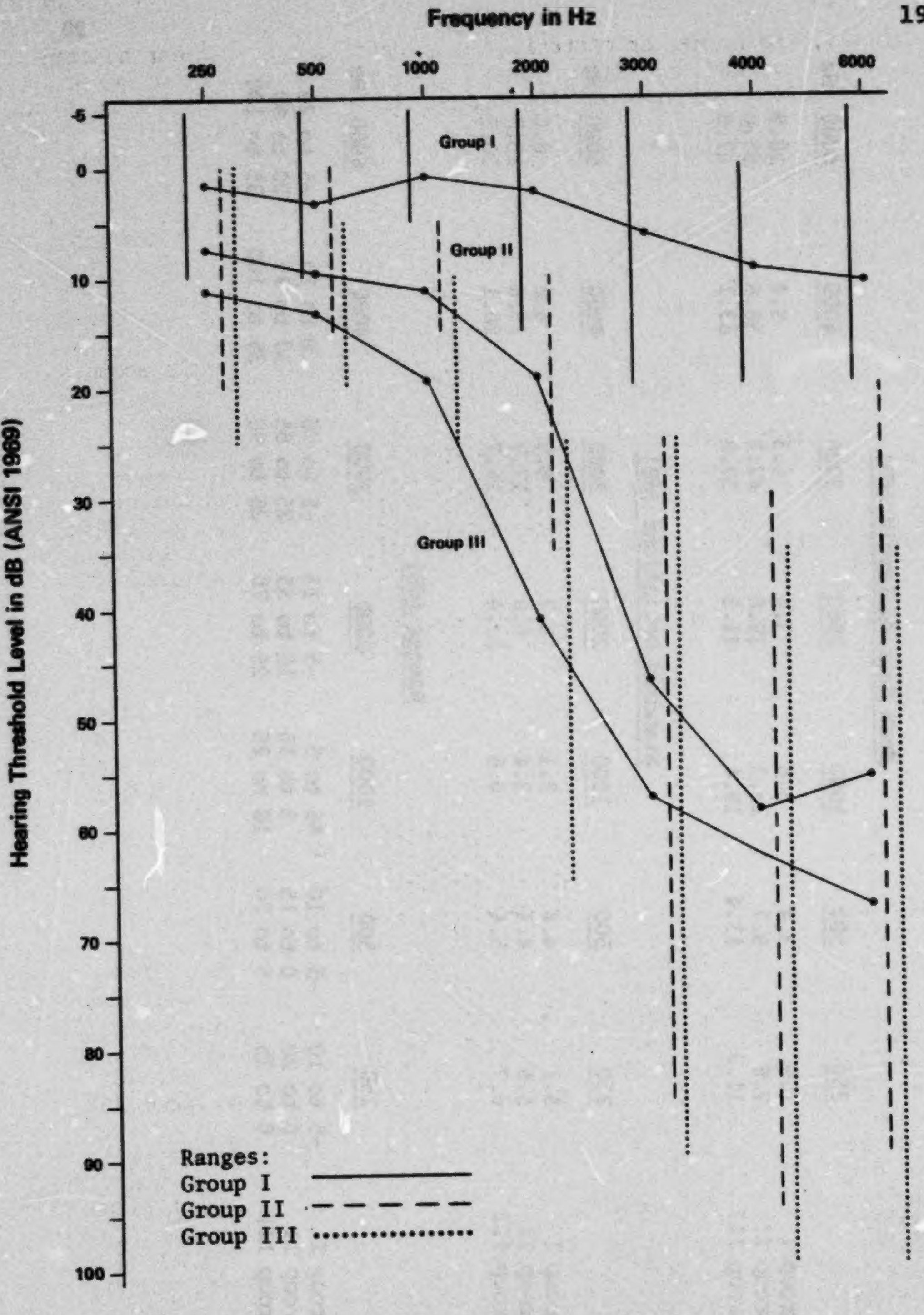

Figure 1. Mean better-ear hearing levels and ranges of the three experimental groups.

Table I. Means, standard deviations, and ranges of better-ear hearing levels in dB for each audiometric frequency.

Mean Hearing Levels (dB)

	250	500	1000	2000	3000	4000	6000 Hz
Group I	1.9	3.4	0.6	3.4	6.3	9.4	10.9
Group II	7.8	9.7	11.3	19.0	47.2	58.8	55.6
Group III	11.3	13.4	19.4	41.3	57.8	63.1	67.8

Standard Deviations (dB)

	250	500	1000	2000	3000	4000	6000 Hz
Group I	5.1	4.4	3.1	5.7	8.5	6.0	8.0
Group II	5.5	4.3	3.4	9.0	17.7	15.6	18.7
Group III	6.7	5.4	4.8	13.4	16.7	20.1	20.3

Ranges (dB)

	250	500	1000	2000	3000	4000	6000 Hz
Group I	-5 to 10	-5 to 10	-5 to 5	-5 to 15	-5 to 20	0 to 20	-5 to 20
Group II	0 to 20	0 to 15	5 to 15	10 to 35	25 to 85	30 to 95	20 to 90
Group III	0 to 25	5 to 20	10 to 25	25 to 70	25 to 90	35 to 100	35 to 100

Table II. Means, standards deviations, and ranges of better-ear hearing levels in dB for various combinations of frequencies.

Mean Average Hearing Levels (dB)

	500,1000, 2000 Hz	500,1000, 2000, 3000 Hz	1000,2000, 3000 Hz	1000,2000, 4000 Hz	2000,3000, 4000 Hz	3000,4000, 6000 Hz
Group I	2.5	3.4	3.4	4.4	6.6	8.8
Group II	13.4	22.0	25.9	29.6	41.6	54.0
Group III	24.7	33.0	39.6	41.2	54.1	62.9

Standard Deviations (dB)

	500,1000, 2000 Hz	500,1000, 2000, 3000 Hz	1000,2000, 3000 Hz	1000,2000, 4000 Hz	2000,3000, 4000 Hz	3000,4000, 6000 Hz
Group I	3.4	4.4	4.9	3.8	5.7	6.4
Group II	2.9	5.6	7.2	5.0	10.3	15.1
Group III	3.1	5.7	8.2	8.4	14.3	17.6

Ranges (dB)

	500,1000, 2000 Hz	500,1000, 2000, 3000 Hz	1000,2000, 3000 Hz	1000,2000, 4000 Hz	2000,3000, 4000 Hz	3000,4000, 6000 Hz
Group I	-3 to 8	-4 to 11	-3 to 13	0 to 13	-2 to 18	-2 to 18
Group II	10 to 18	14 to 30	15 to 35	20 to 37	25 to 63	27 to 90
Group III	20 to 28	21 to 43	25 to 55	28 to 55	30 to 80	37 to 93

Table III. Age and educational level of the three groups.

	N	Age (years)			Educational Level (Years)		
		Mean	SD	Range	Mean	SD	Range
Group I	16	37.8	8.9	21 to 55	13.6	2.2	12 to 18
Group II	16	45.3	7.8	27 to 54	12.6	3.1	9 to 20
Group III	16	50.7	5.4	38 to 56	11.6	2.1	9 to 17

(1975) found decreases in performance as both age and noise level increased. Inspection of Table III shows that age levels of the three groups in the present experiment were not equivalent. There was slightly more than a decade between the ages of groups I and III. However, it was expected that the contribution of aging would be minimal since the speech materials selected were not as difficult as PB words, an age limit was set in the mid-fifties, and there was some attempt to match the groups for aging.

Speech Materials

Both sentence and monosyllabic materials were used in order to provide added means of comparison with other research, and so that the two types of materials could be compared in similar test conditions. One of these was a standardized form of the Revised Central Institute for the Deaf (RCID) Sentences. The recorded form of these sentences is known as the University of Maryland Test #1 (UM Test #1) (Elkins, Causey, Beck, Brewer, and de Moll, 1975). The other was a test of monosyllabic discrimination, the Modified Rhyme Test (MRT). Speech materials are presented in Appendix C.

UM Test #1

Sentence material was selected because of its close relationship to connected discourse (Giolas and Epstein, 1963; Giolas, 1966), and because it is the form of speech material that the AAOO rule was supposedly based on. Webster (1969) has recommended the use of sentences for finding out how everyday speech will be heard, and Davis (1973) has recommended the use of these particular (CID) sentences for validating the AAOO assumptions.

The original CID sentences were developed by researchers at the Central Institute for the Deaf in response to criteria set forth by the Committee on Hearing and Bio-Acoustics (CHABA) of the National Research Council. CHABA's criteria were that the sentences should closely resemble "everyday" speech in such parameters as vocabulary, sentence length, grammatical structure, and redundancy (Silverman and Hirsh, 1955). The sentences that constitute the UM Test #1 were modified from the original lists in order to achieve homogeneity of sentence length by Harris, Haines, Kelsey, and Clack, (1961) to form the Revised CID sentences, (RCID).

The UM Test #1 is a tape recording of the ten RCID lists of ten sentences each, recorded by a male talker with general American speech. The test was standardized on 100 normal-hearing subjects and 55 subjects with various degrees of sensori-neural hearing loss (Elkins et al, 1975). The investigators found a fairly steep performance-intensity function, resembling the curve for spondee words, and that the slope was steeper for hearing-impaired than for normal-hearing subjects. The UM Test #1 was originally recorded in the anechoic chamber of the University of Maryland's

Biocommunications Laboratory. The tape consists of ten lists of ten sentences, each about seven words in length. Each list contains 50 key words for which correct responses are scored. Each sentence is preceded by the carrier phrase "Number ___". The subjects respond by repeating the whole sentence (minus the carrier phrase), or as much of it as is perceived.

The tape of the UM Test #1 was re-recorded in order to juxtapose desired lists, and a speech-shaped noise was recorded at the beginning of the tape in order to serve as a calibration signal. A pilot study was then conducted to examine the equivalency of the lists and to check for a learning effect. (See discussion of Pilot Study #1 in Appendix A.)

MRT

The Modified Rhyme Test is an outgrowth of the Fairbanks Rhyme Test (1958), which was modified by House, Williams, Hecker, and Kryter (1965) in order to provide a multiple choice, closed-set test for the purpose of assessing speech communication systems. It was modified very slightly once more, recorded by three talkers, and mixed with three levels of speech-shaped background noise in order to produce what is known as the Modified Rhyme Hearing Test (MRHT) (Kreul, Nixon, Kryter, Bell, Lang, and Schubert, 1968). The resulting test was intended primarily for clinical purposes. According to Kreul et al it should be "capable of rank-ordering patients according to their ability to discriminate speech under 'everyday' listening conditions." Aside from a few word changes in order to eliminate some objectionable words and to reduce word redundancy, the lists are essentially the same as those of House et al (1965).

There are several reasons why the MRT was selected. First, material less redundant than the CID sentences was considered useful. Kryter (1973) points out that monosyllabic material is sometimes characteristic of everyday speech (the unfamiliar name, the important telephone number, the unexpected message, etc.). Also, the MRT is noted for ease of administration and scoring (Kryter and Whitman, 1965; Kreul et al, 1968), it can be used with untrained listeners, and is reported as providing fairly stable scores (House et al, 1965; Kreul et al, 1968; Nabelek and Pickett, 1974a and 1974b). Additional advantages of the MRT are that it is a closed-set task, eliminating problems of vocabulary familiarity, and that the subject needs only to circle the selected response item. The write-down feature was desirable since the subjects' responses to the other speech materials were oral. Each form contains 50 sets of 6 words each, one of which acts as a test word and the other 5 as foils. Thus, there are 6 different test lists for each form.

For this experiment the Kreul tapes (MRHT) were not selected because the background noise, which had been mixed with the speech signal, would have prevented the use of any other level or type of masking stimulus. Instead, eight lists of the Kreul version of the MRT were recorded in quiet, by a male talker with general American speech and good voice control. The recording procedures and carrier phrase were the same as those described by Kreul et al (1968). The lists were recorded in a sound-treated room, on precision quality tape, using a Sony TC-850 tape recorder with a Sony electret condenser microphone ECM-270. Each test word was imbedded in the carrier phrase, "You will mark the ________, please." Each item was given six seconds: three seconds for the utterance and three seconds' pause. Timing was monitored by an oscilloscope trace that was recycled every six seconds, and voice level was monitored by a VU meter. Seven practice sessions were conducted before the final recordings were made.

All eight lists were employed in a pilot experiment on normal-hearing subjects using a fixed speech-to-noise ratio. The four lists with the closest means and least variability were selected for the final experiment. (See discussion of Pilot Study #2 in Appendix A.) The tape was then copied to include only the four selected lists and one practice list. A speech-shaped noise was recorded at the beginning of the tape to be used as a calibration signal.

Masking Noise

The masking noise used was a "babble" of twelve talkers. A tape consisting of sound-on-sound recordings of six voices, three male and three female, producing a babble of twelve voices, was supplied by Karl Pearsons of Bolt, Beranek, and Newman, Inc. This tape was rerecorded so as to provide a continuous 20-minute segment of babble, and a speech-shaped calibration signal was applied to the beginning of the tape. The range of level fluctuation of the babble was a maximum of 2 dB.

Babble was selected as a masker because of its lifelike quality (subjects reported that it sounded like a party), and because of its speech-like spectrum. A mixture of many voices was first used by Miller (1947) who concluded that "the best place to hide a voice is among other voices." Babble has been used as a masker also by Nabelek and Pickett (1974a and 1974b), and by Miner and Danhauer (1976).

Relationship of Speech and Noise Stimuli

Speech Level

A review of the literature concerning "everyday" speech levels reveals that the frequently-cited levels of 65 to 70 dB (unweighted long-term rms level measured one meter from the talker) are not necessarily typical of conversational speech. Gardner (1966) found that the level of conversational speech in a "free-space" room was approximately 50 dB (B-weighted) and in a quiet office, 58 dBB. When subjects were asked to read prepared text in the same conditions, their voice levels were 6 to 8 dB higher. Pearsons, Bennett, and Fidell (1976) found a similar phenomenon when subjects were asked to recite a memorized passage. Voice levels at one meter in anechoic conditions were an average of 52 dBA for "casual" (conversational) effort and 57 dBA for "normal" effort (reciting prepared material). These levels would be approximately 3 dB higher when measured on the linear scale of a sound level meter (Kryter, 1970). It appears, therefore, that conversational levels in quiet are approximately 52 to 60 dB (long-term rms level), depending on the acoustics of the room and the conditions of the conversation.

Studies by Kryter (1946), Korn (1954), Pickett (1958), Webster and Klumpp (1962), and Gardner (1966) show that individuals automatically raise their voices as background noise levels increase. Korn postulates an increase in speech level of 0.38 dB for every 1-dB increase in noise. Webster and Klumpp found a 0.7 dB/dB increase, while both Kryter and Pickett estimate a 0.3 dB/dB increase. Pearsons et al (1976) also studied voice levels in noise, and made measurements in field as well as laboratory conditions. Speech levels were measured at the listener's ear, which was usually about one meter from the talker. The authors found average speech levels of 57 dBA in urban homes and 55 dBA in suburban homes, with ambient noise levels of 48 dBA and 41 dBA, respectively. They also found speech levels of 61 dBA in department stores, 73 dBA in trains, and 77 dBA in airplanes. By comparing speech levels with background noise levels the authors concluded that individuals raise their voices about 0.6 dBA for every 1-dBA increase in noise, from 48 dBA up to about 67 dBA, at which point talkers and listeners move closer than one meter in order to maintain intelligibility. Accordingly, the following levels of speech in noise could be expected at the listener's ear:

Speech dBA	55	57	63	69
Noise dBA	41	48	58	68

These relationships could become less favorable if communicating distances were greater than one meter, if individuals failed to move closer together at higher levels, or if the talker failed to raise his or her voice.

A speech level of 60 dBA was selected for this study to reflect a slightly raised conversational voice. This level could commonly occur outdoors, inside department stores or inside urban homes (Pearsons et al, 1976).

Speech-to-Noise Ratio

The relationships between speech and noise in this study were determined mainly by the location of the performance-intensity functions of certain normal-hearing and hearing-impaired subjects (see discussion of Pilot Study #3 in Appendix A). Speech-to-noise ratios were selected so that the hearing-impaired subjects achieved discrimination scores between approximately 20% and 80%. The normal-hearing pilot subjects had lesser amounts of difficulty for the same conditions. Slightly different speech-to-noise ratios were employed for the two materials since the performance-intensity function of the UM Test #1 has been reported to be somewhat steeper (Elkins et al, 1975) than that of the MRT (House et al, 1965). As a result of the pilot work, the speech-to-noise ratios selected were 0, -3, and -6 dB for the MRT and -1, -3, and -5 dB for the UM Test #1. A quiet condition for each material was also included in order to provide a basis for comparison. (The residual sound level of the unoccupied room was 19 dBA, so the speech-to-noise ratio for the quiet condition was approximately +40 dB.)

These speech-to-noise ratios are not always typical of "everyday" speech communication. They are not implausible, however, since any of the conditions listed above could occur (conversationalists more than one meter apart, one party fails to raise his voice, etc.). Negative speech-to-noise ratios do occur on aircraft (Pearsons et al, 1976) and undoubtedly occur in other noisy situations such as cocktail parties, city streets, and factories.

Sound Field

The speech materials were presented in a mildly reverberant test space. The reverberation time of the room was 0.625 second, as measured with pink noise, (see section on measurement). This reverberation time was not unlike everyday conditions. Nabelek and Pickett (1974a and 1974b) noted that reverberation times of 0.5 to 1.0 second are typical of small-to-medium sized rooms with hard-wearing, non-porous surfaces (such as school rooms or cafeterias). Reverberation times used in their studies were 0.3 and 0.6 second. In other speech intelligibility studies Harris (1965) used a 5-second reverberation time, Millin (1968) used 0.45 sec., Acton (1970) used 0.3 second, MacKeith and Coles (1971) used 0.35 second, Aniansson (1973) used 0.5 second, and Kuznairz (1973) used 0.2 second. Nabelek and Pickett (1974b), and Bullock (1967), report that hearing-impaired listeners are more sensitive to increased reverberation than normal listeners. Also, subjective reports of hearing-impaired individuals indicate an adverse effect of reverberation on speech discriminiation.

Listening Mode

Although binaural listening more closely resembles real life conditions, a monaural, better ear condition was used for this study. Experimental evidence shows that normal-hearing and even hearing-impaired individuals experience a binaural advantage in noise, especially in reverberant conditions (Moncur and Dirks, 1967; MacKeith and Coles, 1971; Nabelek and Pickett, 1974a and 1974b). Therefore, in assessing the speech discrimination abilities of hearing-impaired subjects, there was the danger that binaural hearing would influence discrimination scores to an unknown and uncontrollable extent.

For this reason a pilot study was conducted on normal-hearing subjects in the proposed experimental conditions. The rationale was that if no binaural advantage were evident for normal listeners, it would be safe to conclude that the hearing-impaired listeners, whose poorer ears would contribute even less than if they had binaurally symmetrical hearing, would not be affected. However, the results showed a binaural advantage of 13% for normal listeners, and therefore the monaural condition was selected (see discussion of Pilot Study #2). In order to occlude the non-test ear, a combination of an earplug and a monaural earmuff was utilized.

Instrumentation

Equipment

All experimental tests were conducted in a double-walled Industrial Acoustics Co. Model 1205-A sound-treated room that met the requirements for background noise in audiometer rooms established by ANSI (1960). Dimensions of the room were 9'10" wide by 9'2" long by 6'6" high. It had been paneled throughout with formica-surfaced particle board in order to create a mildly reverberant space.

Inside the test room subjects were seated in the middle of the room, facing a corner. A two-way window was located to the subjects' right so that they and the investigator could see each other easily. A Grason-Stadler Model 162-4 loudspeaker was located on the floor in the corner, the base of which was approximately 80 inches from the subject's head. A KLH Model Six loudspeaker was located directly above the Grason-Stadler speaker approximately 65 inches from the subject's head. The location of each was marked with masking tape to assure a constant position throughout the experiment.

Signal generating and monitoring equipment was located outside the test room. The recorded noise and speech materials were each reproduced on Kudelski Nagra Type D3 tape recorders, attenuated by Hewlett-Packard 350 C attenuators, and amplified by SWTP 207/A power amplifiers. The speech signal was fed into the KLH loudspeaker and the babble into the Grason-Stadler loudspeaker. A Balentine Model 320 true rms voltmeter was situated so that it could be connected to the output of the tape recorder, attenuator, or amplifier of either the speech or noise system in order to monitor the voltages.

Hearing threshold tests were performed with a Grason-Stadler Model 1701 diagnostic audiometer with TDH-49 earphones in MX/41 AR cushions. Oral responses to the UM Test #1 were picked up by a Shure lavalier-type microphone, routed through the Grason-Stadler audiometer, and recorded by a Sony TC-850 tape recorder. They were also monitored through a small loudspeaker in the Grason-Stadler audiometer. Both of these systems provided adequate intelligibility of the subjects' responses, even at the lowest speech-to-noise ratio. A simplified block diagram of the experimental equipment is shown in Figure 2.

Measurement

Pink noise was used to measure the reverberation of the test room. It was generated by a Hewlett-Packard 8057A precision noise generator, and measured with a Bruel and Kjaer 4145 condenser microphone, and a Bruel and Kjaer Type 2305 Level Recorder. The reverberation time of the room was 0.625 second, measured according to the ANSI, S1.1 (1960) definition: "The reverberation time of a room is the time that would be required for the mean-square sound pressure level therein, originally in a steady state, to decrease 60 dB after the source is stopped."

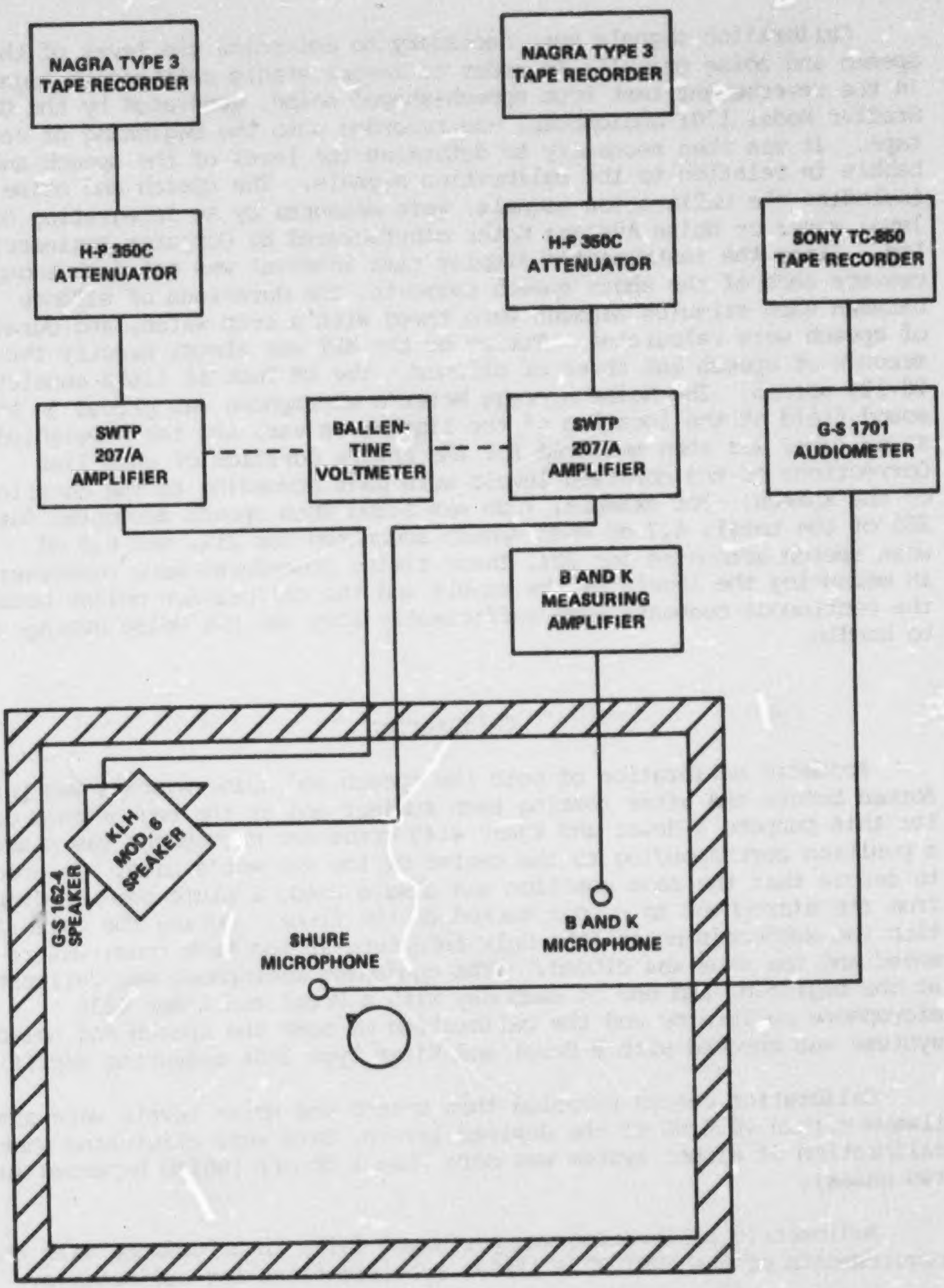

Figure 2. Block diagram of equipment.

Calibration signals were necessary to determine the level of the speech and noise stimuli. In order to insure stable calibration signals in the reverberant test room speech-shaped noise, generated by the Grason-Stadler Model 1701 audiometer, was recorded onto the beginning of each tape. It was then necessary to determine the level of the speech and babble in relation to the calibration signals. The speech and noise, including the calibration signals, were measured by an integrating sound level meter or Noise Average Meter manufactured by Computer Engineering Ltd. Since the instrument's display time interval was not fast enough to measure each of the short speech segments, the durations of silence between each stimulus segment were timed with a stop watch, and durations of speech were calculated. Timing of the MRT was almost exactly three seconds of speech and three of silence. The UM Test #1 lists consisted of 20-22% speech. The Noise Average Meter's microphone was placed in the sound field at the location of the listener's ear, and the A-weighted sound level was then measured for the entire duration of each list. Corrections to the observed levels were made according to the duration of the speech. For example, 7 dB was added when speech accounted for 20% of the total, 6.7 dB when speech accounted for 21%, and 6.5 dB when speech accounted for 22%. These timing procedures were unnecessary in measuring the levels of the babble and the calibration noises because the continuous segments were sufficiently long for the Noise Average Meter to handle.

Calibration

Acoustic calibration of both the speech and noise systems was performed before and after testing each subject and at the end of each day. For this purpose a Bruel and Kjaer 4145 condenser microphone was placed in a position corresponding to the center of the subject's head. In order to assure that the same position was always used, a plumb bob was dropped from the microphone to a spot marked on the floor. During the calibration the subject's chair (the only furniture in the test room) was removed and the door was closed. The condenser microphone was calibrated at the beginning and end of each day with a Bruel and Kjaer 4230 microphone calibrator and the calibration of both the speech and noise systems was checked with a Bruel and Kjaer Type 2606 measuring amplifier.

Calibration checks revealed that speech and noise levels were almost always within $\underline{+}0.5$ dB of the desired levels. Data were eliminated when calibration of either system was more than 1 dB off (which happened in two cases).

Audiometric calibration was performed daily in accordance with the requirements of the ANSI S3.6 (1969) specifications for audiometers.

Hearing Protective Devices

In order to achieve a monaural condition, or at least to minimize any binaural contribution, a V-51R earplug was inserted by the investigator in the subject's poorer ear. Four sizes of the plug were available, and the plugs were fitted according to the procedures recommended by Guild (1966). In order to achieve maximum attenuation a set of David Clark 117 earmuffs was rendered monaural and the remaining muff was placed over the plugged ear. Table IV lists attenuation values for the V-51R earplug, the David Clark 117 earmuffs, and the combination of the two devices. These data represent mean values minus one standard deviation, and were supplied by the U.S. Air Force (1973).

Table IV. Attenuation values in dB for hearing protectors as reported by U.S. Air Force (1973). Data represent mean values minus one standard deviation.

Hearing Protector	Frequency in Hz 125	250	500	1k	2k	3k	4k	6k	8k
David Clark 117	13	21	36	37	33	34	35	33	39
V-51R	22	20	22	24	33	34	29	33	31
Combined	29	36	35	36	38	48	52	46	38

Method of Presentation

All subjects were instructed about the general purpose of the experiment, and the length and nature of the test session. They were then asked to read and sign a standard consent form used by the U.S. Air Force, Aerospace Medical Research Laboratory (See Appendix B.).

Air conduction and bone conduction audiometry were performed according to the modified Hughson-Westlake technique described by Newby (1964). Subjects with differences of more than 10 dB between air conduction and bone conduction thresholds were not used in the study.

The poorer ear was then occluded with an earplug and earmuff, and the subject was seated in the test room. The floor had been marked with masking tape to ensure a constant position for the subject's chair. In order to prevent the subject from turning the unoccluded ear toward the loudspeakers a plumb bob was adjusted to be about two inches in front of the subject's nose. The subject was asked not to touch it, and to glance at it periodically to make sure that it was in the center of his or her line of vision. Standard instructions were then read aloud by the experimenter while the subject read them silently (see Appendix C.) Two practice lists were administered, one for each type of material at speech-to-ratios of -1 and -3 dB. After the practice session the subjects were given a short rest, and then instructed about the remainder of the test. In the middle of the experiment they were given a ten-minute rest between administration of the two speech materials. The total test lasted approximately two hours.

During the main experiment the speech-to-noise ratios were changed by adjusting the attenuator of the noise system. This was done between each list (or pair of UM Test #1 lists) according to a predetermined program for each experimental subject. The order of difficulty (speech-to-noise ratio of -1 or -3) of the practice lists was counterbalanced across subjects, as was the order of presentation of the experimental speech materials. The order of lists within each type of material was also counterbalanced, and the order of speech-to-noise ratio was pseudo-randomized so as to minimize the chance occurrence among subjects of the same speech-to-noise ratio with the same list.

Subjects responded to the MRT by selecting the correct response from one of six alternatives for each test item. They responded to the UM Test #1 by repeating each sentence immediately after it was presented. During the instructions all subjects were encouraged to guess. They were told to complete every item on the MRT, even though they were not sure of the stimulus word. For the UM Test #1 they were told to repeat those words they did hear, even though they might not have heard the entire sentence, and always to guess when unsure.

Responses to the UM Test #1 were monitored by the investigator through the talkback system of the Grason-Stadler audiometer, and were recorded by a Sony TC-850 tape recorder. Later these responses were scored by two other listeners, so as to minimize the possibility of error on the part of a single listener. Since product-moment correlations of the scorings of the investigator with each of the two other listeners were 0.995 and 0.990, the investigator's scorings were used.

Analysis of the Data

Speech discrimination scores in percentage correct were tabulated for all of the experimental data. The data were then subjected to a three-factor analysis of variance using a repeated measures model (Winer, 1971). Simple main effects were examined to clarify certain interactions, and Newman-Kuels tests were performed on pairs of means. The parameters studied were hearing loss (according to groups), speech-to-noise ratio, and type of speech material.

Pearson product-moment correlations were performed in order to test the predictive capabilities of different combinations of frequencies. A correlational matrix was prepared which related hearing loss in the following combinations of frequencies to speech discrimination scores: 500, 1000, and 2000 Hz; 500, 1000, 2000, and 3000 Hz; 1000, 2000 and 3000 Hz; 1000, 2000, and 4000 Hz; 2000, 3000, and 4000 Hz; and 3000, 4000, and 6000 Hz. The results of these analyses are discussed in the next chapter.

Corrections for Guessing

In order to minimize the occurrence of correct responses due to chance the MRT was corrected for guessing as recommended by Kryter (1972) and in the proposed ISO standard for measuring the intelligibility of speech (1975):

$$I \text{ in } \% = \frac{100}{T} \; R - \frac{W}{N-1}$$

where I = intelligibility
T = number of items in the test
N = number of alternatives (6 for MRT)
R = number right
W = number wrong

Such a correction was not considered appropriate for the UM Test #1, since the number of alternatives was so large, and virtually unknown. When this correction had been performed the mean discrimination scores for the two speech materials became somewhat more similar to each other for all conditions except quiet. Table V shows the mean discrimination scores for the MRT before and after correcting for guessing. Mean scores for the UM Test #1 are shown for comparison.

Three-Factor Design

Slightly different speech-to-noise ratios were used for the two speech materials because of the differences between the shape of their performance-intensity functions. Because of these differences it was originally planned that the data for each speech material would be analyzed separately in two two-factor analyses. However, on inspection of the mean data for each speech-to-noise ratio in each group, it was decided that the speech-to-noise ratios could be combined into four levels, thus facilitating the study of speech materials in a three-factor analysis of variance. They were combined as follows:

Speech-to-Noise Ratio (dB)

UM Test #1	Quiet	-1	-3	-5
MRT	Quiet	0	-3	-6
Combined	Quiet	0-1	-3	-5-6

Table V. Mean discrimination scores in percent correct for MRT before and after corrections for guessing, and for UM Test # 1.

	MRT before Correction				MRT after Correction				UM Test #1			
S/N in dB	Q	0	-3	-6	Q	0	-3	-6	Q	-1	-3	-5
Group I	94	86.8	80.1	69.1	92.8	84.1	76.1	62.9	99.6	87.7	78	57.9
Group II	89.1	69	59.4	45	87	62.8	51.3	34	97.8	67.7	50.8	28.6
Group III	82.8	61.1	43.1	32.4	79.3	53.4	31.8	19.7	93.9	53.2	36.8	15.3

S/N = Speech-to-noise ratio
Q = Quiet

CHAPTER IV

RESULTS

This chapter will present the results of the analysis of the data. The outcome of the analysis of variance will be shown, as well as the tests of simple main effects and the Newman-Keuls procedures. Two additional procedures will be presented—a series of correlations between combinations of audiometric frequencies and discrimination scores, and an ad hoc partitioning of the groups according to high-frequency, rather than mid-frequency hearing acuity. Mean scores and standard deviations in terms of percent correct responses are shown in Table VI. Data are given for the three groups of subjects on each of the two speech materials in four speech-to-noise ratios.

Analysis of Variance

The data were subjected to a three-factor analysis of variance using a repeated measures model (Winer, 1971). The three main effects studied were hearing loss, speech-to-noise ratio, and speech materials. There were three levels of hearing loss, represented by groups I, II and III; four levels of speech-to-noise ratio: quiet, 0 and -1 dB combined, -3 dB, and -5 and -6 dB combined; and two speech materials, the UM Test #1 and the MRT. The dependent variable was discrimination score in percent correct responses. The level of significance that was determined for the study was the .05 level of confidence. However, the .01 level was reported when it occurred. The results of the analysis displayed in Table VII show that differences in discrimination scores for the main effects of hearing loss and speech-to-noise ratio were significant at the .01 level of confidence (hearing loss F = 37.9, df 2,45; and speech-to-noise ratio F = 630, df 3,135). The difference in discrimination scores due to the main effect of speech materials was not significant at the .01 level but was significant at the .05 level of confidence (F = 5.47, df 1,45). The interaction of hearing loss and speech-to-noise ratio was significant at the .01 level (F = 25.7, df 6,135), and the interaction between speech materials and speech-to-noise ratio was significant at the .01 level (F = 16.6, df 3,135). The interaction between speech materials and hearing loss, and the three-way interaction between speech materials, hearing loss, and speech-to-noise ratio were not significant.

Since two of the four interactions were statistically significant, tests on the simple main effects of the hearing loss and speech materials factors were performed (according to Winer, 1971) in order to examine their interactions with speech-to-noise ratio. The results of these tests and tests using the Newman-Keuls procedure for assessing the difference between ordered means (Winer, 1971) are displayed in Table VIII. Examination of the hearing loss factor as it interacted with speech-to-noise ratio revealed significant differences among the three

Table VI. Mean scores and standard deviations in percent correct responses.

Group I

MRT*

S/N	Q	0	-3	-6
Mean	92.8	84.1	76.2	63.0
SD	5.0	4.1	7.0	9.4

UM#1

S/N	Q	-1	-3	-5
Mean	99.6	88.2	78.1	58.4
SD	0.5	6.4	8.2	14.3

Group II

MRT*

S/N	Q	0	-3	-6
Mean	87.0	62.8	51.3	34.0
SD	6.3	12.4	14.3	14.0

UM#1

S/N	Q	-1	-3	-5
Mean	97.8	67.7	50.8	28.6
SD	2.6	15.4	18.6	16.4

Group III

MRT*

S/N	Q	0	-3	-6
Mean	79.3	53.4	31.8	19.7
SD	9.1	16.3	18.3	21.1

UM#1

S/N	Q	-1	-3	-5
Mean	93.9	53.2	36.8	15.3
SD	7.0	22.2	20.7	13.4

S/N = Speech-to-noise ratio in dB

*MRT scores have been corrected for guessing.

Table VII. Summary of analysis of variance to determine effects of hearing loss, speech-to-noise ratio and speech materials on speech discrimination scores.

Source of Variation		Sum of Squares	df	Mean Squares	F
	Between Subjects				
Factor A	Hearing Loss	67,473	2	33,737	37.9**
Error A	(Subjects within groups)	40,078	45	891	
	Within Subjects				
Factor B	S/N	156,971	3	52,324	630**
AB	Hearing Loss x S/N	12,793	6	2,132	25.7**
Error B	(B x Subjects within groups)	11,214	135	83	
Factor C	Speech Materials	738	1	738	5.47*
AC	Hearing Loss x Speech Materials	55	2	27.5	.20
Error C	(C x Subjects within groups)	6,054	45	135	
BC	S/N x Speech Materials	2,913	3	971	16.6**
ABC	Hearing Loss x S/N x Speech Materials	435	6	72.5	1.24
Error BC	(BC x Subjects within groups)	7,897	135	58.5	

S/N = Speech-to-noise ratio

* Significant at .05 level of confidence.

** Significant at .01 level of confidence.

Table VIII. Tests on significant interactions of hearing loss with speech-to-noise ratio, and speech materials with speech-to-noise ratio.

Simple Main Effects

Hearing loss		
Quiet	df 2,180	F = 7.93 **
S/N 0-1 dB	df 2,180	F = 32.10 **
S/N -3 dB	df 2,180	F = 53.14 **
S/N -5-6 dB	df 2,180	F = 55.78 **
Speech materials		
Quiet	df 1,180	F = 35.9 **
S/N 0-1 dB	df 1,180	F = 2.63
S/N -3 dB	df 1,180	F = 1.51
S/N -5-6 dB	df 1,180	F = 7.03 **

Newman - Keuls

Quiet		
Difference in % correct	I and II	3.8
between means of groups	I and III	9.6 **
	II and III	5.8 *
S/N 0-1 dB		
Difference in % correct	I and II	20.6 **
between means of groups	I and III	32.6 **
	II and III	12 **
S/N -3 dB		
Difference in % correct	I and II	26 **
between means of groups	I and III	42.7 **
	II and III	16.7 **
S/N -5-6 dB		
Difference in % correct	I and II	29.1 **
between means of groups	I and III	42.9 **
	II and III	13.8 **

S/N = Speech-to-noise ratio
* Significant at the .05 level of confidence.
** Significant at the .01 level of confidence.

groups, even in the quiet condition. These differences were significant at the .01 level of confidence for all four speech-to-noise ratios and they became more pronounced as speech-to-noise ratios decreased. Further probing using the Newman-Keuls technique showed that the means of all three groups were significantly different from each other at the .01 level of confidence in the three noise conditions. In the quiet condition groups I and III were significantly different at the .01 level, also groups II and III at the .05 level, but not groups I and II. The interaction between hearing loss and speech-to-noise ratio is graphically displayed in Figure 3. Since the increment in noise level between quiet and the first noise condition is much greater than those between the three noise conditions the curves have been broken between quiet and the speech-to-noise ratio of 0-1 dB.

Examination of the speech materials factor as it interacted with speech-to-noise ratio revealed differences significant at the .01 level of confidence between the two materials in the quiet condition with higher scores on the UM Test #1, and in the most difficult noise condition with higher scores on the MRT. Since only two means were involved the Newman-Keuls procedure was not performed. The interaction between the speech materials and speech-to-noise ratio factors is graphically displayed in Figure 4. Because neither the interaction between hearing loss and speech materials, nor the three-way interaction between hearing loss, speech-to-noise ratio, and speech materials was significant, additional tests were not performed.

Correlational Tests

In order to determine which frequencies, or groups of frequencies, best predicted speech discrimination scores, groups II and III were combined. Each subject's audiogram was divided into various frequency combinations (500, 1000, and 2000 Hz; 500, 1000, 2000, and 3000 Hz; 1000, 2000, and 3000 Hz; 2000, 3000, and 4000 Hz; 3000, 4000, and 6000 Hz). For each subject each of these combinations was correlated with the subject's speech discrimination scores (averaged across the two speech materials) for the four experimental conditions, and with a composite discrimination score obtained by averaging across the two materials and the four conditions. Pearson product-moment correlations were calculated, which are displayed in a correlational matrix in Table IX. Most of the different audiometric combinations show high positive correlations with each other, which is to be expected. All of the combinations show negative correlations that are significantly different from zero with the discrimination scores resulting from all of the conditions tested, indicating that the greater the hearing loss, the lower the discrimination score. Those combinations that include frequencies above 2000 Hz show particularly high correlations with discrimination scores, especially in the higher levels of background noise.

Since all of the frequency combinations showed correlations with discrimination scores that were significantly different from zero, it was necessary to explore the relationships of the correlations to each

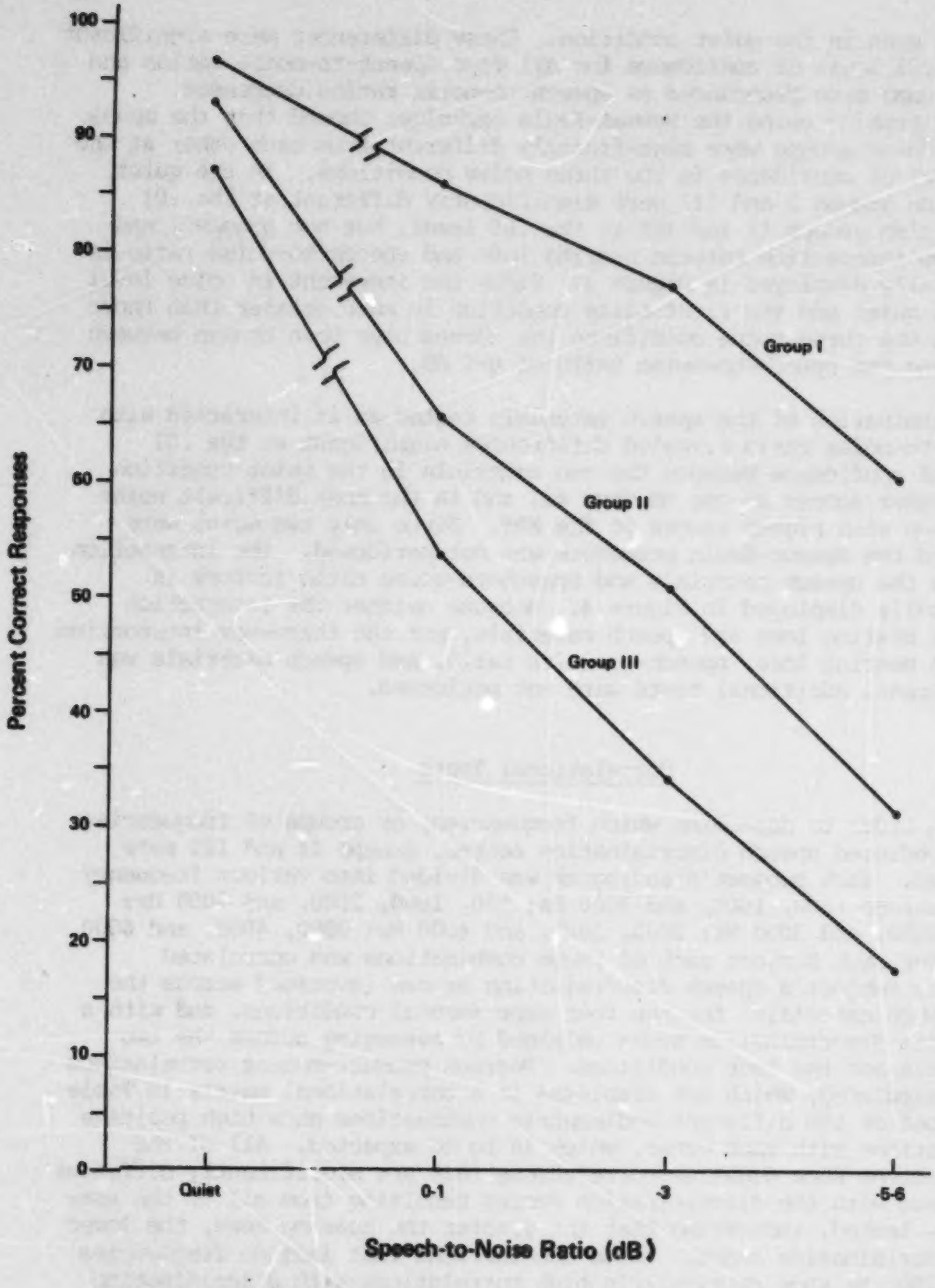

Figure 3. Mean percent correct responses of the three groups as a function of speech-to-noise ratio. Scores are averaged across the two speech materials.

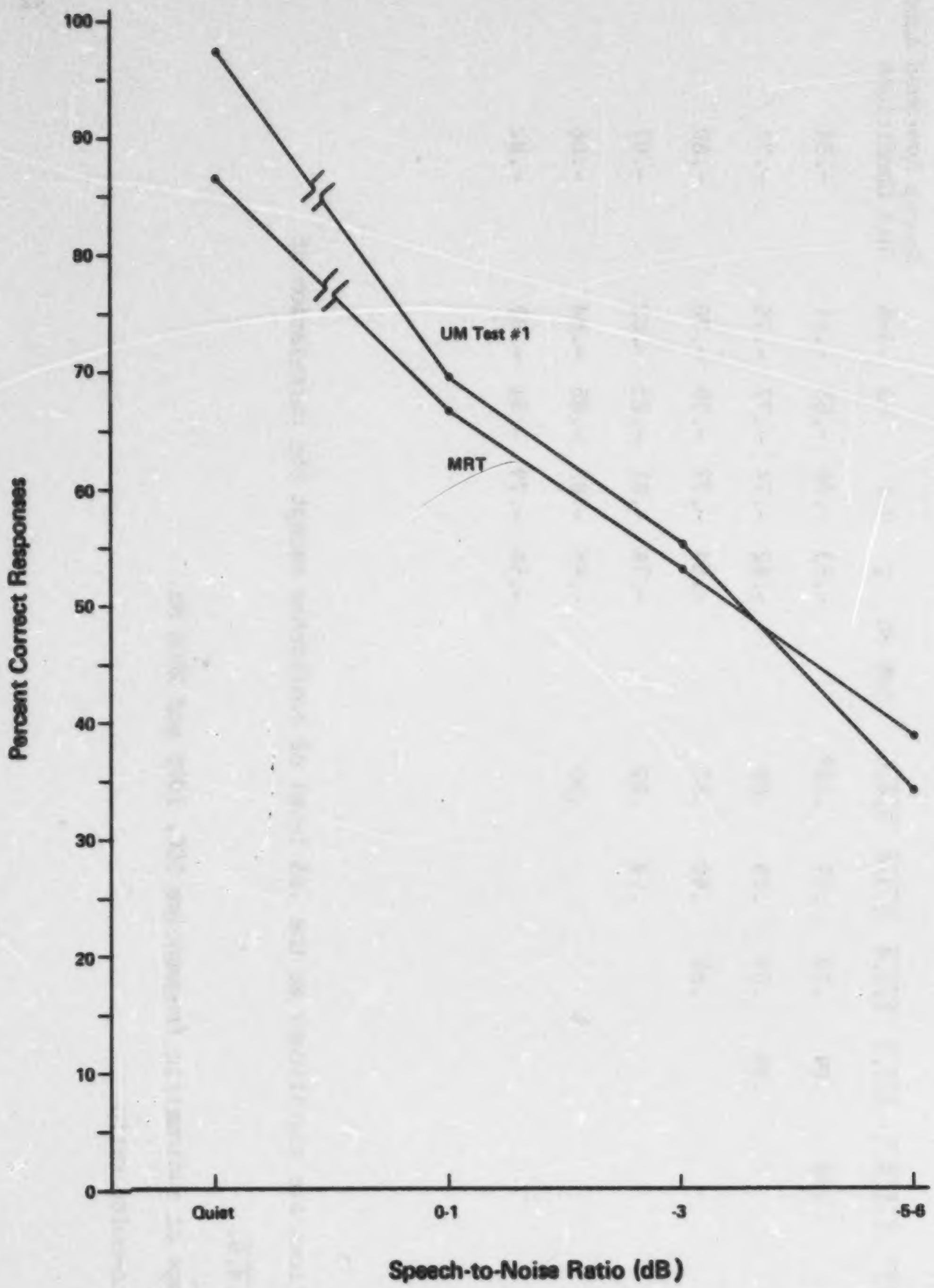

Figure 4. Mean percent correct response in the two speech materials as a function of speech-to-noise ratio. Scores are averaged across the three groups.

Table IX. Correlational matrix among six combinations of audiometric frequencies and discrimination scores in various speech-to-noise ratios. Data are for subjects in groups II and III.

	$\overline{.5,1,2}$	$\overline{.5,1,2,3}$	$\overline{1,2,3}$	$\overline{1,2,4}$	$\overline{2,3,4}$	$\overline{3,4,6}$	S/N dB	Q	0-1	-3	-5-6	Scores Averaged Across All Conditions
$\overline{.5,1,2}$		.86	.84	.73	.57	.32+		-.53	-.48	-.55	-.51	-.54
$\overline{.5,1,2,3}$			.99	.89	.89	.69		-.62	-.74	-.77	-.75	-.77
$\overline{1,2,3}$				.89	.90	.67		-.64	-.77	-.79	-.76	-.80
$\overline{1,2,4}$					.94	.82		-.76	-.81	-.85	-.82	-.87
$\overline{2,3,4}$						.90		-.65	-.83	-.85	-.84	-.86
$\overline{3,4,6}$								-.56	-.79	-.81	-.82	-.82

+ All correlations are significant at the .05 level of confidence except the correlation of $\overline{.5,1,2}$ with $\overline{3,4,6}$.

$\overline{.5,1,2}$ = Average of audiometric frequencies 500, 1000 and 2000 Hz.

S/N = Speech-to-noise ratio

Q = Quiet

other. The Hotelling formula (Guilford, 1965) was used in order to test the significance of differences between the correlations. The outcome is in the form of a t-ratio that takes into account the correlation of frequency combinations with each other, as well as the correlation of frequency combinations with discrimination scores. The matrix in Table X displays the significance of differences among correlations of frequency combinations with the composite discrimination scores (the last correlation displayed in Table IX). According to the Hotelling procedure significant differences exist between the 500, 1000, and 2000 Hz combination and all of the other combinations employing higher frequencies. The other frequency combinations are not significantly different from each other with the exception of 500, 1000, 2000, and 3000 Hz in comparison with 1000, 2000, and 4000 Hz. It can be concluded that for the present experimental conditions, the average of 500, 1000, and 2000 Hz is the poorest predictor of speech discrimination, while the other combinations are about equally efficient, with 1000, 2000, and 4000 Hz appearing to be slightly superior in quiet, and 2000, 3000, and 4000 Hz a better predictor in noise.

Further Ad Hoc Comparisons

During the experiment it had become evident that there were some members of Group II whose discrimination scores were poorer than certain members of Group III. High-frequency hearing acuity appeared to be the critical factor. Therefore groups II and III were divided on the basis of high-frequency, rather than mid-frequency hearing levels to see if the differences in discrimination scores increased.

Groups II and III were combined and then partitioned according to whether subjects' thresholds were better or worse than the median hearing level (47 dB) for the averaged frequencies 2000, 3000, and 4000 Hz. Those subjects whose average hearing levels fell on the median were eliminated, leaving a total number of 29 subjects, 14 of whom had hearing levels better than median (Group Y) and 15 worse (Group Z). Mean discrimination scores of these two groups as a function of speech-to-noise ratio are shown as solid lines in Figure 5. Mean scores for groups I, II, and III are shown by dotted lines for comparison. The differences between the newly partitioned groups are greater than when the groups are divided by average hearing level at 500, 1000, and 2000 Hz.

The causes and implications of the results presented above will be discussed in Chapter V.

Table X. Matrix displaying significance of differences among correlations of various frequency combinations and speech discrimination scores averaged over the four experimental conditions.

	$\overline{.5,1,2}$	$\overline{.5,1,2,3}$	$\overline{1,2,3}$	$\overline{1,2,4}$	$\overline{2,3,4}$	$\overline{3,4,6}$
$\overline{.5,1,2}$		3.96**	4.51**	5.11**	3.67**	2.63*
$\overline{.5,1,2,3}$			1.97	2.33*	2.03	.69
$\overline{1,2,3}$				1.64	1.43	.29
$\overline{1,2,4}$					.33	.98
$\overline{2,3,4}$						.96

* Significant at the .05 level of confidence.

** Significant at the .01 level of confidence.

$\overline{.5,1,2}$ = Average of audiometric frequencies 500, 1000, and 2000 Hz (etc.)

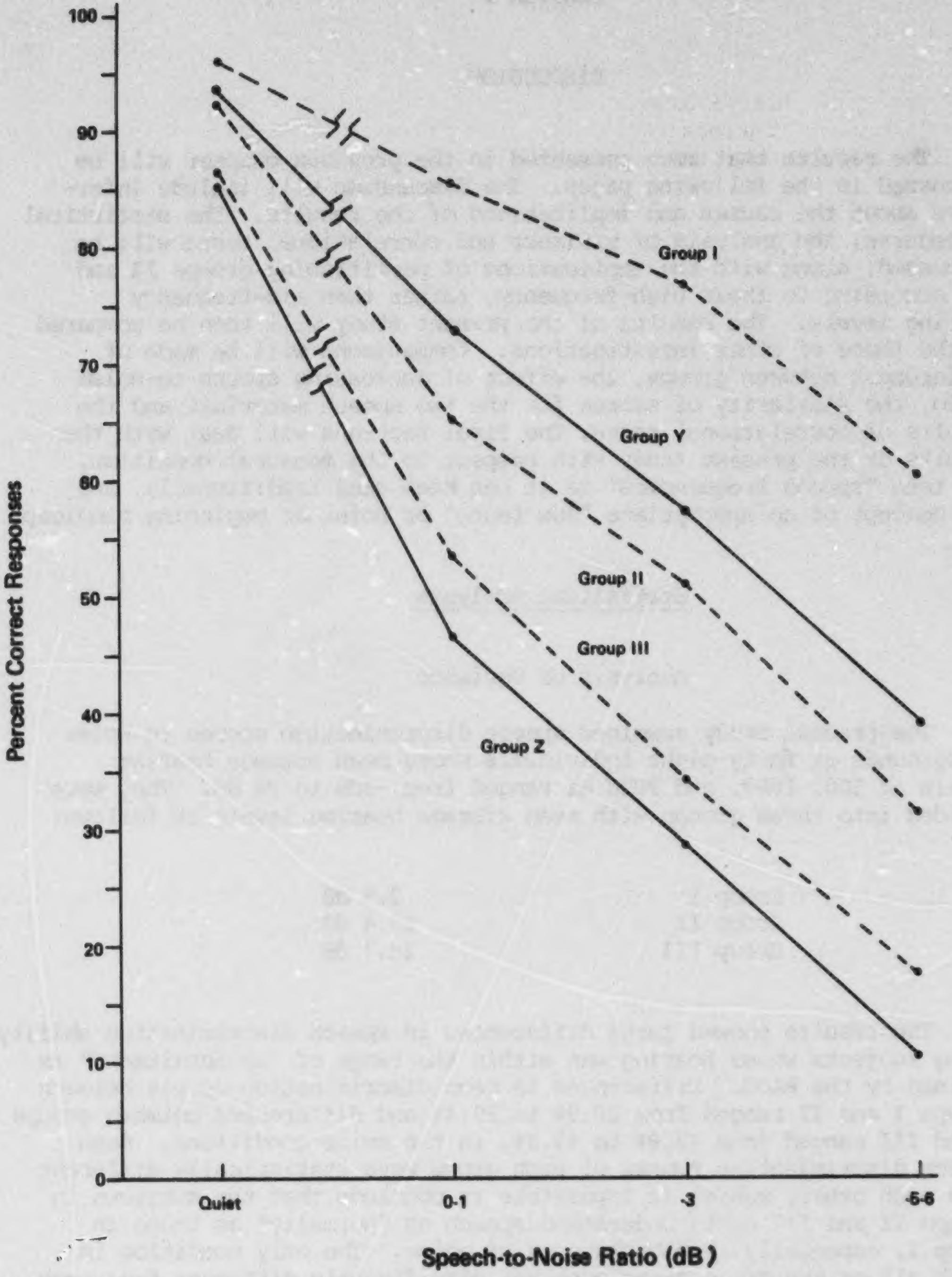

Figure 5. Mean speech discrimination scores in percent correct as a function of speech-to-noise ratio. Groups II and III have been combined and partitioned according to whether subjects have more or less hearing loss than the median (47 dB) at the average of 2000, 3000, and 4000 Hz. Group Y = <47 dB, Group Z = >47 dB.

CHAPTER V

DISCUSSION

The results that were presented in the previous chapter will be discussed in the following pages. The discussion will include inferences about the causes and implications of the results. The statistical procedures, the analysis of variance and correlational tests will be discussed, along with the implications of partitioning groups II and III according to their high-frequency, rather than mid-frequency hearing levels. The results of the present study will then be compared to the those of other investigations. Comparisons will be made of differences between groups, the effect of decreasing speech-to-noise ratio, the similarity of scores for the two speech materials and the results of correlational tests. The final sections will deal with the results of the present study with respect to the monaural condition, the term "speech frequencies" as it has been used traditionally, and the concept of an appropriate "low fence" or point of beginning handicap.

Statistical Analysis

Analysis of Variance

The present study examined speech discrimination scores in noise backgrounds of forty-eight individuals whose mean average hearing levels at 500, 1000, and 2000 Hz ranged from -3dB to 28 dB. They were divided into three groups with mean average hearing levels as follows:

Group I	2.5 dB
Group II	13.4 dB
Group III	24.7 dB

The results showed large differences in speech discrimination ability among subjects whose hearing was within the range of "no impairment" as defined by the AAOO. Differences in mean discrimination scores between groups I and II ranged from 20.9% to 29.4% and differences between groups I and III ranged from 32.9% to 43.2%, in the noise conditions. Mean speech discrimination scores of each group were statistically different from each other, making it impossible to conclude that the subjects in groups II and III could understand speech as "normally" as those in Group I, especially in a background of noise. The only condition in which all of the group means were not significantly different from each other was quiet, where the difference between mean scores of groups I and II was not statistically significant. However, even in quiet groups I and II, and II and III were significantly different. These findings indicate that differences between individuals traditionally categorized as unimpaired may appear even under listening conditions that are favorable. The differences between groups increased markedly as listening conditions become less favorable, indicating that the introduction of noise exacerbated the effects of hearing loss.

Differences between groups and between speech materials may have been underestimated due to truncation effects in the quiet and -5-6 dB conditions. Eighteen subjects scored 100% correct on the sentences in the quiet condition, and five subjects (all from Group III) scored 0% on the UM Test #1, four of whom scored 0% on the MRT, in the -5-6 dB condition.

Differences between discrimination scores for the two speech materials were fairly small, especially after the scores on the MRT had been corrected for guessing. However, the differences were significant at the .01 level of confidence in quiet and in the most difficult noise condition, with subjects scoring higher on the UM Test #1 in the former, and lower on the UM Test #1 in the latter condition. Since the interaction between hearing loss and speech materials was not significant, it is assumed that the differences between groups were not dependent upon speech materials, and the differences between speech materials were not a function of hearing loss. Also, since the three-way interaction between hearing loss, speech materials and speech-to-noise ratio was not significant, it is assumed that neither the hearing loss nor the speech materials factors were dependent upon an interaction of the other variable with speech-to-noise ratio.

The fact that the two speech materials were presented at slightly different speech-to-noise ratios does not explain the differences in mean scores of the two materials in the -5-6 dB condition. This procedure probably decreased rather than increased the difference in the -5-6 dB condition, and may have masked a difference in the 0-1 dB condition. Since the MRT was presented at a speech-to-noise ratio of 0 dB, it can be assumed that the scores would have been slightly poorer at -1 dB, where the UM Test #1 was presented. Likewise, since the UM Test #1 was presented at a speech-to-noise ratio of -5 dB, it can be assumed that the scores would have been slightly poorer if it had been presented at -6 dB as was the MRT.

These results imply that the two materials could be considered equivalent only in the -3 dB condition. However, the lack of interaction between the hearing loss and speech materials factors indicates that the two materials did not differ in their ability to elicit differences between the three groups.

Other Tests on the Data

Pearson product-moment correlations were calculated to assess the abilities of various frequency-averaging methods to predict speech discrimination scores. The combinations of 1000, 2000, and 4000 Hz and 2000, 3000, and 4000 Hz showed the highest correlation with discrimination ability (or more rightly disability, since the correlations were negative).

It appeared that the average of 1000, 2000, and 4000 Hz was a slightly better predictor in the quiet condition, while the average of 2000, 3000, and 4000 Hz was slightly better in the noise conditions. All of the combinations that included frequencies above 2000 Hz showed high correlations with speech discrimination that were not significantly different from each other, with the exception of 1000, 2000, and 4000 Hz, which was a significantly better predictor than 500, 1000, 2000, and 3000 Hz. However, all of the combinations tested were significantly better than 500, 1000, and 2000 Hz. At least for the present experimental population and conditions the combinations of 1000, 2000, and 4000 Hz and 2000, 3000, and 4000 Hz are slightly superior to the others, especially for predicting speech discrimination scores in a background of noise. It is reasonable to conclude that these frequency combinations are superior to the traditionally-used 500, 1000, and 2000 Hz for describing a population with high-frequency hearing loss.

In order to assess further the contribution of high-frequency hearing acuity groups II and III were combined, and then partitioned according to hearing levels that were greater or less than the median hearing level (47 dB) at the averaged frequencies 2000, 3000, and 4000 Hz. This procedure revealed differences of approximately 30 percentage points between the two groups in all but the quiet condition. These differences are much greater than those between the original groups II and III, divided according to average hearing level at 500, 1000, and 2000 Hz. The magnitude of these differences suggests that a more appropriate method of grouping the hearing-impaired subjects would have been according to hearing level in the higher frequencies, rather than the method followed, which was based on average hearing levels at 500, 1000, and 2000 Hz.

Comparison of Results with Other Studies

Differences Among Groups

The division of hearing-impaired subjects into groups resembles the studies of Acton (1970) and Kuzniarz (1973). The designs of the three studies are roughly analogous, with some exceptions. Acton and Kuzniarz each used three hearing-impaired groups instead of the present study's two. The two more severely-impaired groups had hearing levels that were slightly greater than the present study's groups II and III respectively. Other differences are that both Acton and Kuzniarz used PB monosyllables instead of the MRT, they did not employ sentences (although Kuzniarz did use sentences in another part of his experiment), and Kuzniarz' speech materials were in Polish. In spite of these differences, the overall results of the present study are quite consistent with those of Acton and Kuzniarz in most respects. Clear differences were evidenced between hearing-impaired and normal listeners, and among the various experimental groups. The only exception was Acton's Group A, whose hearing levels were between those of Group I and Group II of the present study. Acton's Group A performed nearly as well as his control group in speech-to-noise ratios of +10 dB and above, and they surpassed the control group in less favorable speech-to-noise ratios. The author attributes the phenomenon to "conditioning" that may have occurred in Group A's experience in noisy industrial environments.

The scores of all subjects in the present experiment are slightly higher than those of the other two investigators, which is somewhat surprising since the monaural condition was employed in the present study but not in the others. The most likely explanation is the fact that the other investigators used PB monosyllables instead of the MRT and sentences. According to Kryter and Whitman (1965), the difference between PB words with a large number of alternatives (1000) and the MRT is about 20 percent at a speech-to-noise ratio of 0 dB. The ANSI S3.5 (1969) standard for calculation of the Articulation Index (AI) shows a difference of 25 to 30 percent between 1000 Harvard PB words and rhyme words and sentences, respectively at an AI of 0.4, which is a speech-to-noise ratio of approximately 0 dB in the present experiment. Similarity between Acton's PBs (Fry's monosyllabic word lists), Kuzniarz' PBs (in Polish), and the Harvard word lists, can only be assumed since the author knows of no comparative data. Other explanations for the differences in scores between the present study and those of Acton and Kuzniarz could be in the slight differences in hearing levels between analogous groups, possible differences in the methods of measuring the speech signals, and in the masking capabilities of the different noises, (although the spectra of the three noises appear to be fairly similar). Aside from the differences in scores, certain common conclusions can be drawn for the three studies, namely that there are clear differences between the discrimination scores of individuals that traditionally have been considered either mildly impaired or unimpaired and those with truly normal hearing (as defined by these investigators), and that high frequency hearing acuity plays an integral part in defining these differences.

Effect of Speech-to-Noise Ratio

Other investigations as well as the present one have shown increasing differences between normal-hearing and hearing-impaired groups as speech-to-noise ratio decreases. The scores of Acton's (1970) hearing-impaired groups moved further apart as speech-to-noise ratio decreased from +20 dB toward 0 dB, but the control group's scores remained between experimental groups A and B. Kuzniarz (1973) found increasing differences between groups as the speech-to-noise ratio grew more difficult, and also Dickman (1974) found that differences between the normal and hearing-impaired groups increased as they progressed from "easy" to "difficult" speech-to-noise ratios for different types of noises.

Both Ross et al (1965) and Elkins (1971) found that hearing-impaired individuals performed more poorly than normal-hearing subjects in noise backgrounds, but did not find greater differences as noise levels increased. In fact, for certain conditions the relative shift for hearing-impaired subjects was smaller than it was for normals. The most likely explanation for this phenomenon is that these investigators presented the stimuli at 40 dB above each subject's speech reception threshold, whereas the other investigators (Acton, Kuzniarz, Dickman, and the present author) presented the stimuli at the same level for all subjects. (See discussion of the effects of this procedure in Chapter II).

Another explanation may involve the introduction of reverberation. According to Sabine (1950), the effect of reverberation is to cause an overlapping or blurring of speech segments by their predecessors in time. This effect is achieved by the reflection of sound by hard surfaces. The faster the sound decays (or the shorter the reverberation time) the less likely is this blurring effect to occur. Further complicating the situation is the fact that the signal consists of both direct and reflected components. Bullock (1967), and Nabelek and Pickett (1974a) have shown that hearing-impaired listeners are more sensitive to increased reverberation than are normal listeners. Nabelek and Pickett (1974a) hypothesized on the basis of their data that hearing-impaired listeners are not as efficient as their normal-hearing counterparts at integrating delayed speech reflections with the direct sound. It is possible that increasing the level of the babble in relation to the speech signal produced a greater degree of reflected speech-like sound, which in turn increased the difficulty of hearing-impaired subjects in selecting the desired signal from the unwanted background.

Similarity of Speech Materials

The fact that there were only small differences between the two types of speech materials is not surprising. The ANSI S3.5 (1969) standard for calculating the Articulation Index contains curves for the two materials that are very similar. In the ANSI figure sentence scores are about 5% higher than scores on rhyme tests for more favorable AI values. The curves overlap at an AI of about 0.3 and below this point scores on the rhyme test are slightly higher for comparable AI values. The relationship between the two materials (shown in Figure 4) is comparable to the relationship between the curves shown in the ANSI standard.

Comparison with Other Correlational Studies

When a variety of frequency averages were correlated with speech discrimination scores, the results showed the importance of high-frequency hearing. The averages of 1000, 2000, and 4000 Hz and 2000, 3000, and 4000 Hz appeared to be the most efficient predictors of speech discrimination, and all of the combinations that included frequencies above 2000 Hz showed significantly higher correlations than the average of 500, 1000, and 2000 Hz.

These findings are not consistent with some of the earlier correlational studies, such as those of Carhart (1946), Harris et al (1956) and Quiggle et al (1957). However, as mentioned in Chapter II, these studies were conducted in quiet conditions. Also, "hearing for speech" was defined in terms of speech reception thresholds rather than discrimination scores. Carhart (1946) and Quiggle et al (1957) compared pure-tone hearing acuity to speech reception thresholds for spondee words. It is not surprising that the lower frequencies (500, 1000, and 1500 Hz) were the ones that correlated most highly with speech reception thresholds, since spondee words are primarily dependent upon vowels and minimally dependent upon consonants for intelligibility. Also of interest is the fact that Quiggle et al (1957) based their conclusions on random samples drawn from the Wisconsin State Fair of 1954, rather than from hearing-impaired populations. Harris et al (1956) developed a multiple regression formula that was based on speech reception thresholds for PB words, (the level at which 50% were identified correctly), and therefore consonant energy was a more important factor than it was with spondee tests. Interestingly, the investigators found that the most important frequencies were 1000, 2000, 4000, 500, and 6000 Hz in that order. The resulting regression formula was found to be the best of a variety of formulae for predicting hearing loss for speech in patients with sloping losses (20 dB or more difference between 500 and 2000 Hz). But the best predictor for all types of hearing loss was determined to be an adjusted version of the 500, 1000, and 2000 Hz average.

The present study's findings are more consistent with the results of other investigators, when speech discrimination score rather than speech reception threshold was the dependent variable. Mullins and Bangs (1957), although they used quiet instead of a noise background, found that 2000 and 3000 Hz were the audiometric frequencies that best predicted speech discrimination scores. Harris (1965) found that the frequency region of 2000-4000 Hz was the most important for understanding distorted speech (without noise in the background). However, since everyday speech is not always distorted, the authors concluded that the average of 1500, 2000, and 3000 Hz would be the best predictor of discrimination scores when the speech was distorted about 50% of the time. Harris' study was criticized by Webster et al (1965) for distorting the material so heavily, and then for presenting it at 40 dB above the subjects' speech reception thresholds. These procedures might respectively increase or decrease the importance of high-frequency hearing in correlational analyses.

Results of correlations performed on data that had been gathered in backgrounds of noise have been inconclusive. Ross et al (1965) found that 500 Hz correlated significantly with "relative discrimination shift" in noise, but poorly with speech discrimination in quiet, while the opposite was true of 4000 Hz. The authors suggested that one reason for this unexpected finding may have been the fact that the hearing-impaired subjects already had poor discrimination scores in quiet, and therefore the introduction of masking did not produce a very large relative discrimination shift. Another explanation may be the fact that the experimenters presented the stimuli at a level of

40 dB above speech reception threshold, the implications of which practice have been discussed earlier.

These explanations may also apply to the studies by Elkins (1971), Murry and Lacroix (1972), and Myers and Angermeier (1972), all of which used the Modified Rhyme Hearing Test (MRHT) and presented the stimuli at 40 dB above speech reception threshold. In Elkins' study the relative discrimination shift for hearing-impaired listeners was smaller than expected. Correlations of all audiometric frequencies with speech discrimination in noise were fairly low, but the individual frequencies 2000, 3000, and 4000 Hz did correlate significantly with speech discrimination in the least noisy conditions. The results of Murry and Lacroix were similar to those of Elkins. They found that correlations between the average of 1000, 2000, and 3000 Hz and discrimination scores "approached reliability" for the easy lists of the MRHT, but that correlations for the noisier conditions were uniformly low. Myers and Angermeier found large amounts of scatter when discrimination scores were plotted as a function of audiometric frequency. (Correlations were not computed.) The authors concluded that it was impossible to explain the variance in scores by any audiometric index.

The results of the present study are in agreement with two other correlational studies of speech discrimination in noise. Lindeman (1971) tested discrimination of monosyllables (Dutch) in "cocktail party" noise. The author found that the best predictors of speech discrimination in noise were the audiometric frequencies 2000 and 6300 Hz, respectively. Kryter et al (1962) found that the frequencies 2000, 3000, and 4000 Hz were consistently good predictors of speech discrimination scores both in high noise (-3 dB S/N) and low noise (+10 dB S/N) conditions. Although correlations were higher for PB words than for sentences, these frequencies were also superior to the others) (500, 1000, and 6000 Hz) for predicting sentence discrimination. Multiple correlations for various combinations of frequencies showed that the combination of 1000, 2000, and 3000 Hz was only slightly less efficient than the combination of 2000, 3000, and 4000 Hz. The authors concluded that, in light of earlier studies that advocated greater importance of the lower frequencies, the combined frequencies 1000, 2000, and 3000 Hz would be a good compromise.

Webster (1964) criticized the study of Kryter et al (1962) for the application of correlational techniques to a population with predominantly high-frequency hearing loss. Webster pointed out that correlation coefficients would be influenced by the range and number of cases distributed throughout the range of measurement. Since the population in question showed greater ranges and numbers of cases with losses above 2000 Hz than below 2000 Hz, Webster maintained that the correlations of discrimination scores with higher frequencies were artificially high. This criticism would also apply to the present study since there was considerably more variability and a wider range of thresholds in the higher audiometric frequencies than in the lower ones. However, it would be inappropriate to perform these tests on a population with similar ranges and variability in low-frequency as well as high-frequency thresholds, if the results were to apply to individuals with sensori-neural hearing losses. In spite of Webster's implication to the contrary in 1964, the vast majority of today's compensation cases involve sensori-neural hearing losses, nearly all of which are noise-induced. To include subjects with significant losses at 500 Hz would necessitate changing the character of the population by including a substantial number of conductive losses. Thus, the population would no longer be suitable for studying the formula which is supposed to describe it. Webster does admit that if Kryter's recommendation (of a 1000, 2000, and 3000 Hz method) is confined to noise-induced-hearing loss subjects, then it is "optimal for the sample to be studied."

Monaurality

The monaural condition was used in this experiment, even though it was not representative of common, everyday conditions. However, monaurality was needed to accurately assess the relationship of audiometric threshold to speech discrimination scores. The only way to have avoided monaurality would have been to select subjects with completely symmetrical hearing losses binaurally (not a very common condition), which was not possible within the practical constraints of this experiment. Pilot Study #2 showed a binaural advantage of 13% for normal listeners. Therefore, hearing-impaired listeners could conceivably have scored from 0 to 13% more poorly than the normal-hearing subjects if the binaural condition were used, simply because of the difference in thresholds between the two ears. Since the exact amount of influence that this disparity would cause would be unknown, the naturalness of the binaural condition was sacrificed in favor of accuracy. The use of the monaural condition does not diminish the significance of the differences between experimental groups. In fact, such differences would most likely be even larger in the binaural condition because of the relatively smaller contribution of the hearing-impaired subjects' poorer ears.

However, when making comparisons of these data with binaural data collected in a reverberant environment it should be kept in mind that differences between the monaural and binaural data of up to about 13 percentage points could occur, and also that the discrimination scores of all subjects in this experiment could be expected to be somewhat higher in real-life, binaural listening conditions.

The Speech Frequencies

Results of the statistical analyses presented in Chapter IV indicate that the traditional label of "speech frequencies," applied to 500, 1000, and 2000 Hz, is inappropriate. Actually, the term has been applied in a variety of ways in past years. Sabine (1942) referred to all the frequencies between 128 and 4096 Hz as the "important speech range of frequencies" or just the "speech range." The 1947 AMA standard (Carter, 1947) used the frequencies 512, 1024, 2048, and 4096 Hz, and the weighted percentages of hearing loss were "based on the existing data bearing on the relative importance of the auditory frequency and intensity range in the hearing of speech." Even the Committee on Conservation of Hearing of the AAOO (DeForest and Lierle, 1955) referred to these four frequencies as the "speech frequencies."

Fletcher (1950) probably had considerable influence on the use of the term. He found that the frequencies 500, 1000, and 2000 Hz bore the closest relationship to "hearing loss for speech," which was defined as 50% correct responses. As mentioned above, Carhart (1946), Harris et al (1956), and Quiggle et al (1957), defined "hearing for speech" similarly. Davis (1970) supported the practice by referring to the average of 500, 1000, and 2000 Hz as the "central speech range". However, certain investigators have continued to include frequencies above 2000 Hz in the term. For example, Kryter et al (1962) stated that "information in the speech range above 2000 cps contributes significantly to the understandability of sentences in the presence of noise...." Myers and Angermeier (1972) referred to the "speech range" as 500 to 3000 Hz. On the basis of his research, Kuzniarz (1973) objected to "the present concept of so-called 'most important speech frequencies': 500-2000 Hz," and recommended the average of 1000, 2000, and 4000 Hz, which had been accepted by the Ministry of Health in Poland.

The term "speech frequencies" is equated with the average of 500, 1000, and 2000 Hz in most audiology clinics, since this average hearing level corresponds so well with the speech reception threshold level for spondee words (Newby, 1964). It must be remembered, however, that higher frequencies should be included in the definition when predicting speech discrimination ability under everyday conditions. Perhaps in order to avoid confusion the term "mid-frequencies" could be applied to the average of 500, 1000, and 2000 Hz.

Fences

The results of the present investigation have not resolved the question of the location of the point of beginning handicap, or the "low fence". It is evident that the subjects in Group III, whose mean average hearing level at 500, 1000, and 2000 Hz was 25 dB, had considerable difficulty understanding speech in the backgrounds of noise used in this experiment, up to 43 percentage points more difficulty than Group I. This difficulty was encountered for sentences and monosyllables alike. But Group II, whose mean average hearing level at 500, 1000, and 2000 Hz was only about 13 dB also experienced considerable difficulty understanding speech in noise, in this case up to 30 percentage points more difficulty than Group I. Figure 5 showed that a more effective way of dividing the hearing-impaired groups involves higher frequency hearing acuity. This fact, along with the discussion of correlations above, would indicate that whatever fence is selected should include frequencies above 2000 Hz.

In a discussion of the fence issue, Kryter (1970) presented speech intelligibility functions for individuals with various average hearing levels at 500, 1000, and 2000 Hz and at 1000, 2000, and 3000 Hz. The functions were drawn from calculations based on the Articulation Index, where the amount of audible sound pressure in each speech band was determined according to the typical configuration of a noise-induced hearing loss.

Using these AI calculations and speech intelligibility predictions, Kryter devised a method of calculating the percentage of "hearing impairment for speech" for persons with sensori-neural or conductive losses. To account for the added discrimination difficulties that accompany noise-induced hearing loss, even when speech is at an optimal loudness level, Kryter proposed a correction which assigned twice as much handicap to these individuals as compared to those with conductive losses. He also presented different percentages of handicap based on different levels of unamplified speech ("everyday" speech at 65 dB, "conversational" speech at 55 dB, and "weak conversational" speech at 50 dB measured at the listener's ear). For these calculations a fence of 0 dB at the average of 500, 1000, and 2000 Hz or 10 dB at the average of 1000, 2000, and 3000 Hz is implied. In this respect Kryter's proposed method resembles Fletcher's (1929) original "Point-Eight Rule," except that the percentage loss per dB is quite different, the slope being much more steep in Kryter's method.

Assuming that a long-term rms level of 65 dB reflected the level of "everyday" speech, Kryter determined that the hearing level at which individuals could hear 100% of sentences was 15 dB for the average of 500, 1000, and 2000 Hz or 25 dB for the average of 1000, 2000, 3000 Hz. These levels were based on the assumptions that everyday speech environments are quiet (even though the level of 65 dB was chosen to reflect a slightly-raised voice due to ambient noise), that everyday speech is undistorted, and that individuals with noise-induced hearing losses have normally shaped performance-intensity functions (as do individuals with conductive losses). In order to account for lower voice levels, Kryter (1973) proposed as a fence average hearing levels of 6 dB at 500, 1000, and 2000 Hz or 16 dB at 1000, 2000, and 3000 Hz. These levels were consistent with 100% intelligibility of sentences at a level of 55 dB in quiet, but Kryter stated that even these individuals would be disadvantaged in comparison with normal-hearing persons when the speech was distorted, or the level was weaker than usual.

The above proposals have not been accepted by the medical community or by governmental bodies, and the extent to which they are being considered seriously is not known. Most of the fences that Kryter proposed are lower than those presently in use by the various States or Federal agencies.

One way to approach the problem of an appropriate fence would be to find the hearing level at which hearing-impaired subjects begin to perform differently from their normal-hearing controls. A significant finding in Acton's (1970) study was the fact that the most mildly-impaired group (Group A) performed more efficiently than the control group in most of the noise backgrounds, and virtually as well in quiet. By way of explanation the author suggested that in situations where the redundancy of speech is reduced by noise, individuals are able to "get-by" so long as their hearing levels do not reach those of a "critical hearing loss." Once this critical level is passed individuals show increasing difficulty in understanding speech, especially as redundancy is further reduced. Acton hypothesized that the critical hearing loss was somewhere between Group A and Group B. However, both Acton's Group B and this study's Group II performed significantly more poorly than their normal-hearing counterparts in all of the noise conditions, and judging by the magnitude of the differences could be considered past the point of "critical hearing loss."

Mean hearing levels of Acton's groups A and B, and the present study's Group II are as follows:

	$\overline{.5, 1, 2k}$ Hz	$\overline{1, 2, 3k}$ Hz	$\overline{1, 2, 4k}$ Hz
Acton's Group A	4 dB	11.5 dB	14.8 dB
Acton's Group B	13.2 dB	27.6 dB	30 dB
Present Study Group II	13.4 dB	25.9 dB	29.6 dB

Acton's hearing levels have been converted from British Standard 2497 (1954) to ANSI (1969) reference values.*

It could be hypothesized that the appropriate fence lies somewhere between Acton's Group A and the present study's Group II (whose mean thresholds are almost identical with Acton's Group B). If the midpoints between these two groups are selected, then the estimated fences would be approximately 9 dB at 500, 1000, and 2000 Hz, 19 dB at 1000, 2000, and 3000 Hz, and 22 dB at 1000, 2000, and 4000 Hz. These values are very close to the mean minus one standard deviation of Group II hearing levels, which are 10.5 dB, 18.7 dB and 24.6 dB, respectively.

The selection of a fence is ultimately dependent upon the definition of hearing handicap and the conditions under which handicap is assessed. Davis (1965) defined handicap as "the disadvantage imposed by an impairment sufficient to affect one's personal efficiency in the activities of daily living." Since speech communication is the activity most likely to be impaired by hearing loss, the AAOO defined the absence of hearing handicap (or "impairment" in 1959) as, "The ability to hear sentences and repeat them correctly in a quiet environment....," (Lierle, 1959). As mentioned previously, many investigators have pointed out that undistorted speech in quiet is not typical of everyday conditions. Therefore, the determination of hearing handicap should be based on speech discrimination in noise, even if tests in quiet are included.

---

* Audiometric thresholds given in personal communication from Dr. W. I. Acton.

Basically, the selection of a fence is a social issue. It rests on the question of how much speech communication ability is needed in order to conduct the activities of daily living in a satisfactory manner. The answer will undoubtedly be influenced by such variables as an individual's age, occupation, lifestyle and personal preference. Field, rather than laboratory research will probably be needed in order to solve the problem, but research in this area has been inconclusive to date. Until more information is forthcoming, the decision on an appropriate fence will necessarily be somewhat arbitrary.

CHAPTER VI

SUMMARY AND CONCLUSIONS

Summary

The purpose of the investigation was to explore the relationship between hearing level at various audiometric frequencies and speech discrimination in different noise backgrounds. Although these relationships had been investigated numerous times before, the studies had not been designed specifically to test the adequacy of the American Academy of Ophthalmology and Otolaryngology's (AAOO) selection of 26 dB as the "low fence". This hearing level, averaged over the audiometric frequencies of 500, 1000, and 2000 Hz, is said to be the point above which hearing handicap occurs. It has been incorporated into many state compensation statutes, and has also been widely used in the U.S. and abroad for purposes of damage-risk criteria and the setting of occupational noise standards.

The AAOO method for computing hearing handicap has been brought into question during the past few years, both by researchers and by policy-makers, for two primary reasons: a) that the 26-dB fence is too high, and b) for the exclusion of frequencies above 2000 Hz. The present study, therefore, has investigated the relationship between hearing level and speech discrimination to see if there are differences among individuals whose hearing is at or better than the low fence, or whether they are all indeed "not handicapped."

The AAOO low fence is based on the assumption that "hearing impairment should be evaluated in terms of ability to hear everyday speech under everyday conditions." Consequently, the present experiment has employed "everyday" sentences as well as a closed-set test of monosyllables, presented in a quiet background and in various levels of noise, in a mildly reverberant sound field. In designing the study the following experimental questions were posed:

1. What is the relationship between average hearing level at 500, 1000, and 2000 Hz and speech discrimination scores in noise for individuals whose average hearing levels are at or better than the AAOO low fence?

2. Is the relationship dependent upon speech-to-noise ratio?

3. Is the relationship between average hearing level and speech discrimination scores differently described by different speech materials?

4. Which combination of audiometric frequencies best predicts speech discrimination scores?

Forty-eight subjects between the ages of 21 and 56 were tested with two types of speech materials: the University of Maryland Test #1 (UM Test #1), which employs the CID "everyday" sentences, and the Modified Rhyme Test (MRT), a closed-set test of rhyming monosyllables. Subjects were divided into three equal groups according to their better-ear average hearing levels at 500, 1000, and 2000 Hz. Group I had mean average hearing levels of 2.5 dB and hearing at all frequencies (250 - 6000 Hz) of 20 dB or better. Group II had mean average hearing levels of 13.4 dB and Group III had mean average hearing levels of 24.7 dB at 500, 1000, and 2000 Hz. Both groups II and III were unrestricted for hearing loss in the higher frequencies. Since most of these subjects had been exposed to noise their losses were considerably greater in the frequencies above 2000 Hz than in the mid-frequencies.

Each subject listened to the UM Test #1 in a quiet condition and in speech-to-noise ratios of -1, -3, and -5 dB, and to the MRT in quiet and in speech-to-noise ratios of 0, -3, and -6 dB. The reverberation time of the room was 0.625 second and remained unchanged.

Results of the tests and the statistical analyses provided answers to the above questions as follows:

1. Significant differences were found in mean speech discrimination scores among all of the three groups, showing that within the "normal" area under the AAOO fence there was considerable individual variability in the ability to discriminate speech in noise. Groups with mean average hearing levels of 24.7 dB and even 13.4 dB performed significantly more poorly than the control group, whose mean average hearing level was 2.5 dB.

2. The relationship between average hearing level and speech discrimination scores proved to be dependent upon speech-to-noise ratio. The discrimination scores of all three groups were depressed as the speech-to-noise ratio became lower, and the differences between groups increased. These differences were apparent even in the quiet condition, although they became much larger as noise was introduced.

3. Examination of the mean scores of the UM Test #1 and the MRT (after correcting for guessing) showed that subjects scored very similarly on the two kinds of materials in the two intermediate noise conditions, but not in the quiet condition, or in the most difficult noise condition. Mean scores in quiet were generally lower on the MRT than on the UM Test #1, and scores in the most difficult noise condition were lower on the UM Test #1 then on the MRT. The two materials appeared to be equally effective at delineating differences among the three groups in the various noise conditions.

4. Correlational tests revealed that frequency combinations that included frequencies above 2000 Hz were significantly better predictors of speech discrimination scores than the combination of 500, 1000, and 2000 Hz.

Conclusions

On the basis of this experiment it is possible to draw the following conclusions:

1. Individuals whose hearing levels are at or better than the AAOO low fence may have considerably more difficulty in understanding speech than those whose hearing is normal, as defined in this experiment. This is true even for those whose average hearing levels in the mid-frequencies (500, 1000, and 2000 Hz) are approximately 14 dB.

2. Increased levels of noise, in relation to the speech signal, tend to exacerbate the adverse effects of hearing loss.

3. Simple sentences, such as the CID "everyday" sentences, and a closed-set test of monosyllables, such as the MRT, can be considered roughly equivalent for measuring speech discrimination in certain conditions of noise, but not in quiet.

4. Combinations that include frequencies above 2000 Hz are significantly better predictors of speech discrimination score than the combination of 500, 1000, and 2000 Hz for persons with noise-induced hearing loss.

Recommendations

Four principle recommendations can be made as a result of this investigation. First, frequencies above 2000 Hz should be included in any technique for assessing the ability of hearing-impaired individuals to understand speech in "everyday" listening conditions. For the assessment of hearing handicap in a noise-exposed population similar to that of this experiment, the average of 1000, 2000 and 4000 Hz appears to be the most appropriate simple average, since this average has been shown to correlate highly with speech discrimination in quiet as well as in noise. Unequal weighting of the different frequencies was not considered in this report, although it would be reasonable to explore this method in the future using these or other data. It is also recommended that the term "speech frequencies" should not be applied to 500, 1000 and 2000 Hz alone, but should be used broadly to include all of the audible frequencies of the speech spectrum (through 8000 or 10000 Hz), and that frequently-used combinations of audiometric frequencies be specified, such as ".5, 1 and 2k Hz" or "1, 2 and 4k Hz". The combination of 500, 1000, and 2000 Hz could be termed the "mid-frequencies".

The second recommendation pertains to the height of the fence, or the point of beginning handicap. The present 26-dB fence, averaged over 500, 1000 and 2000 Hz has been shown in this investigation to be above the point of beginning handicap. Even 26 dB averaged over 1000, 2000 and 3000 Hz appears to be too high since that level corresponds to the mean hearing level of this study's Group II (at those frequencies), who showed significantly more difficulty in understanding speech than the normal-hearing group. Data gathered in another study (Acton, 1970) showed that individuals with average hearing levels of 12 dB at 1000, 2000 and 3000 Hz, or 15 dB at 1000, 2000 and 4000 Hz performed as well as their normal-hearing controls, even in noise conditions. Therefore, until further research defines this point more precisely, it is suggested that the midpoint between the "handicapped" and "not handicapped" groups is selected, namely, 19 dB at 1000, 2000 and 3000 Hz or 22 dB at 1000, 2000 and 4000 Hz.

The third recommendation is for further research into the concept of the fence as a social issue. Techniques should be developed to determine the amount of speech communication ability that is needed in order to conduct the activities of daily living in a satisfactory manner. Various lifestyles and various activities should be studies. This kind of research would benefit the development of speech communication criteria for normal-hearing as well as hearing-impaired individuals.

A final recommendation pertains to clinical as well as laboratory tests of speech discrimination in noise. In order to assess accurately an individual's ability to understand speech in various "everyday" conditions, speech materials should be presented at a level that reflects life-like listening conditions. The speech level should be the same for all clinical patients or experimental subjects rather than being adjusted to an optimal level for each listener. It appears that the adjustment of presentation level for speech discrimination material to each patient's optimal listening level (PB max) is primarily intended to assess the patient's ability to understand speech with a hearing aid. While this procedure is definitely useful for the intended purpose, it is not appropriate for assessing the speech discrimination abilities of persons who are not suited for or do not intend to wear hearing aids. Instead, individuals should be tested at speech levels of about 38 dB above speech audiometric zero (or 58 dB long-term rms , the speech levels that they typically must listen to in everyday life.

61a

APPENDIX A

PILOT EXPERIMENTS

Instrumentation

Pilot Studies #1 and #2 were conducted about four months earlier than Pilot Study #3, and the placement of equipment was slightly different. For the earlier tests, two of each type of loudspeaker were used, located in corners of the room, and equidistant from the subject at 45 degree angles to the front of the subjects' midline. The noise signal (babble) was routed through KLH Model Six speakers, which were placed on the floor, and the speech signal was routed through Grason-Stadler 162-4 speakers, located directly above the others. Later, for the purpose of creating a more evenly diffuse sound field, one of each type of loudspeaker was removed. For Pilot Study #3 and for the final experiment the subject was seated facing the two remaining loudspeakers, which were placed in one corner of the room. It is believed that this difference in instrumentation does not affect the applicability of Pilot Studies #1 and #2 to the final experiment.

Pilot Study #1

The purpose of this experiment was to examine the equivalency of the UM Test #1 lists in a noise background, and to see whether or not a practice effect occurred, (an increase in scores as a function of familiarity with the task). The ten lists were presented to ten normalhearing subjects in a background of 12-speaker babble. The lists were presented in the sound field and both ears were unoccluded. The level of the speech signal was 55 dB (long-term rms measured at a point corresponding to the listener's ear) and the babble was 62 dB, resulting in a speech-to-noise ratio of -7 dB. Order of presentation of the lists was informally counterbalanced. Subjects' responses were monitored by the investigator through the talkback system of the Grason-Stadler 1701 audiometer.

The resulting speech discrimination scores varied considerably among subjects and among lists for the same subject, but there appeared to be no practice effect. The range of mean scores for the various lists was approximately 20% (not including List A, the recording of which had become defective and subsequently was eliminated). Because of the intra-subject variability and the differences between means of lists, it was decided to pair them on the basis of mean scores for use in the final experiment. This would provide 20 sentences, including 100 key words, for each listening condition. Pairing of the lists brought the mean scores to within 1-1/2% of each other for the speech-to-noise ratio tested (see Table A-1). The selected pairs were lists B + F, D + I, G + H and J + E.

Pilot Study #2

The purposes of this study were to assess potential differences between monaural and binaural discrimination of MRT words by subjects

Table A-1. Pilot Study #1. Mean scores and standard deviations in percent correct of ten normal-hearing subjects for various single and combined lists of the UM Test #1. Speech-to-noise ratio = -7 dB.

List	Mean	SD
B	75.5	11.7
C	67.4	14.1
D	66.7	14.6
E	59.6	7.1
F	55.6	14.8
G	75.6	11.2
H	55.4	11.8
I	67.3	12.7
J	72.8	10.5
Selected Pairs		
B + F	65.6	
D + I	67.0	
G + H	65.5	
J + E	66.2	

with normal hearing, and to determine the equivalency of the MRT word lists as recorded for this experiment.

Eight MRT word lists were recorded, six from MRHT Form 3, and two from Form 2, using the vocabulary and techniques described by Kreul et al (1968). The eight lists were presented to thirteen normal-hearing listeners at a speech-to-noise ratio of -7 dB (speech level 55 dB long-term rms and noise level 62 dB measured at the listener's ear as before). Four lists were presented monaurally and four binaurally and the order of presentation of lists was counterbalanced across subjects and listening mode. Monaurality was achieved by the use of a V-51R earplug and a David Clark type 117 earmuff over one ear whose combined attenuation capabilities were described in the section on measurement. Although the subjects were asked not to move their heads, they were not physically restrained. In order to minimize head movements a plumb line was hung from the ceiling with the bob centered about two inches in front of the head at eye level. Subjects were asked not to touch the plumb bob and to keep it, to the extent possible, centered in the line of vision.

The results showed a mean increase of 13% for binaural over monaural scores (see Table A-2). Consequently, the speech was presented monaurally in the final experiment. Those lists with the smallest standard deviations and the most similar means were selected for use in the main study. Lists B, C, D and F from Form 3 were the best candidates. List A of Form 2 was used as a practice list.

Pilot Study #3

The purpose of this phase of the pilot work was to determine appropriate speech-to-noise ratios so that the scores of the hearing-impaired listeners, and to a lesser extent those of the normal-hearing listeners, would be likely to fall along the linear portion of the performance-intensity function. The subjects consisted of four normal-hearing and four hearing-impaired individuals, (see Table VI). Two of the hearing-impaired subjects would be categorized as Group II (subjects #5 and 6) and two as Group III members (subjects #7 and 8), according to the criteria described earlier. With speech stimuli presented at a fixed level of 60 dBA (measured at the listener's ear), the following speech-to-noise ratios were explored:

	Speech-to-noise ratio (dB)												
MRT	Quiet	+2		0	-1	-2	-3	-4	-5	-6		-8	-9
UM Test #1	Quiet	+2	+1		-1		-3	-4	-5	-6	-7		-9

The four selected MRT lists and the four pairs of UM Test #1 lists were presented monaurally, preceded by a practice list of each type delivered at a mildly difficult speech-to-noise ratio. The data are given in Table A-3. In this table the MRT scores have been corrected

Table A-2. Pilot Study #2. Mean scores and standard deviations in percent correct on the Modified Rhyme Test of thirteen normal-hearing subjects in monaural and binaural listening modes. Speech-to-noise ratio = -7dB. Scores are not corrected for guessing.

MRT		Monaural			Binaural		
List	Form	N	Mean	SD	N	Mean	SD
A	3	7	56.6	12.5	6	69.7	13.6
* B	3	7	55.7	8.8	6	70.0	4.7
* C	3	7	51.7	8.4	6	69.3	10.6
* D	3	7	53.1	6.0	6	67.0	6.8
E	3	6	49.3	13.0	7	62.9	9.6
* F	3	6	57.7	8.0	7	64.9	5.6
A	2	6	64.7	10.2	7	76.0	9.9
B	2	6	61.3	7.9	7	74.3	7.8
			$\overline{X}$ = 56.3			$\overline{X}$ = 69.3	

*Lists selected for final experiment

Table A-3. Pilot Study #3. Individual scores on MRT and UM Test #1 in a variety of speech-to-noise ratios. MRT scores have been corrected for guessing.

		Subject #							
		1	2	3	4	5	6	7	8
		Average hearing level at 500, 1000 and 2000 Hz							
	Speech-to-noise-Ratio (dB)	-5 dB	-2 dB	2 dB	3 dB	12 dB	15 dB	20 dB	28 dB
MRT	Quiet				95.2		95.2	90.4	
	+2		80.8			68.8		61.6	
	+1								
	0			92.8	80.8		83.2		
	-1	80.8	83.2			64		40	64
	-2			92.8					
	-3				78.4		61.6		
	-4	80.8						37.6	40
	-5		66.4	80.8	52	35.2	42.4		
	-6	68.8							44.8
	-7								
	-8			54.4					16
	-9	66.4	30.4			6.4			
UM Test #1	Quiet				99		99	96	
	+2							85	
	+1			98					
	0								
	-1	90		94	93	63	70	71	57
	-2								
	-3	82	84	83	84	41	48		
	-4							19	50
	-5		70	71	54	35	29		
	-6					4			47
	-7	53	54						30
	-8								
	-9	25	30						

for guessing. It appeared that the linear portion of the hearing-impaired subjects' performance-intensity functions was between speech-to-noise ratios of approximately +2 and -7 dB for the MRT and +1 and -5 dB for the UM Test #1. It also appeared that the normal-hearing subjects would have relatively little difficulty with these speech-to-noise ratios. In order to make sure that some degradation would occur in the performance of the normal-hearing subjects, and to highlight differences between groups should such differences occur, the positive speech-to-noise ratios were eliminated and speech-to-noise ratios of 0, -3, and -6 dB were chosen for the MRT, and -1, -3, and -5 dB for the UM Test #1.

67a

APPENDIX B

CONSENT OF VOLUNTEER

DEPARTMENT OF THE AIR FORCE
6570TH AEROSPACE MEDICAL RESEARCH LABORATORY (AFSC)
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

REPLY TO
ATTN OF:

SUBJECT: Consent of Volunteer

TO:

1. I hereby volunteer to participate as a test/experimental subject in the following investigation/test which has as its purpose

____

2.____ has discussed with me to my satisfaction the reasons for this investigation/test and its possible adverse and beneficial consequences.

3. This consent is voluntary and has been given under circumstances in which I can exercise free power of choice. I have been informed that I may at any time revoke my consent and withdraw from the experiment without prejudice and that the investigator or physician may terminate the experiment at any time regardless of my wishes.

4. I understand that before my use as a test subject, I must inform the principal investigator and/or project physician of any change to my medical status. This information will include any medications I have taken and any medical or dental care/treatment received since my last use as a test subject.

____
(Signature of Volunteer)

____
(Project Officer)

____
(Witness)

AMERICAN REVOLUTION BICENTENNIAL 1776-1976

APPENDIX C

INSTRUCTIONS TO SUBJECTS

(MRT)

SPEECH INTELLIBILITY TEST

INSTRUCTIONS

During the test it is important that you do not move your head around. You may move it up and down a little so that you are free to write, but please do not move it sideways. The plumb bob has been arranged so that it would be in the center of your head. Try not to touch it with your head, and as much as possible, keep it in the center of your line of vision.

You will hear some words and sentences in a background of noise. There are 3 parts to the test. The first part is fairly short. First an announcer will read a list of words. Each time the announcer will say "Number one)(or so); you will mark the (word) please." You are to look at the square corresponding to that number, decide which word he said, and circle the word. Try to ignore the noise, even if it sounds rather loud, and concentrate on the voice reading the list of words. If you are not sure which word was said, take a guess.

After that you will hear a list of short sentences. This time you will just repeat each sentence. Please be sure to talk clearly so that I will be able to hear exactly what you have said. If you are not able to hear all the words in the sentence, repeat the ones you do hear. If you are in doubt about some of the words, take a guess.

Now you will hear a series of word lists, and you will circle the test words. Turn the page each time the announcer completes the list of 50 words. Try to ignore the noise and concentrate on the voice reading the words.

After that you will have a few minutes to relax and then you will hear a series of sentences.

During the tests there may be a short pause while I rewind the tapes.

Remember, in all cases if you're not sure of the words you hear, please guess whenever you can.

APPENDIX D

SPEECH MATERIALS

MODIFIED RHYME TEST

NAME____________________________ EAR________________ DATE__________________

MODIFIED RHYME HEARING TEST 3 LIST__________________

1
fang bang
rang hang
gang sang

2
mark bark
park hark
lark dark

3
peel keel
feel eel
reel heel

4
tang tab
tam tap
tack tan

5
sick sit
sing sin
sill sip

6
mass map
mad man
mat math

7
pup pug
putt puff
pun pub

8
hop pop
top cop
shop mop

9
best west
nest rest
test vest

10
cuff cup
cud cut
cub cuss

11
sale sake
safe save
sane same

12
dust rust
just gust
bust must

13
heave heal
heath heap
hear heat

14
dim din
did dig
dip dill

15
took look
cook hook
book shook

16
sap sat
sag sass
sack sad

17
gun run
bun nun
sun fun

18
page pale
pane pay
pave pace

19
got hot
tot pot
lot not

20
tick wick
pick sick
kick lick

21
wit fit
sit hit
bit kit

22
kith kit
kiss kid
king kill

23
foil oil
coil toil
soil boil

24
fig rig
pig wig
big jig

25
peach peas
peal peak
peat peace

26
pill pip
pig pin
pit pick

27
sup sung
sun sum
sud sub

28
fizz fit
fill fib
fig fin

29
bent tent
went dent
sent rent

30
pat pang
pass pan
pad path

31
teach tear
teak team
teal tease

32
dud dun
dub dull
dug duck

33
beak beam
beat bead
beach bean

34
way say
may day
gay pay

35
then hen
pen men
ten den

36
paw say
thaw law
jaw raw

37
lane lace
lake lay
lame late

38
pale tale
bale gale
male sale

39
till bill
fill kill
hill will

40
bed wed
fed led
red shed

41
hold gold
fold cold
sold told

42
bun buff
bug buck
but bus

43
seed seem
seep seen
seethe seek

44
sin tin
win din
fin pin

45
neat heat
beat meat
seat feat

46
fame name
came same
game tame

47
sip rip
hip tip
lip dip

48
bath back
ban bad
bass bat

49
cake cape
case cane
cave came

50
race rate
rake ray
raze rave

UNIVERISTY OF MARYLAND TEST #1

LISTS B THROUGH J

LIST 1B

1. The water's too cold for swimming.

2. Why should I get up so early.

3. Shine your own shoes this time.

4. It's raining right here in the room.

5. Where are you going this morning?

6. You should come here when I call.

7. Don't try to get out of it.

8. We let little children go to the movies.

9. There isn't enough paint to finish.

10. Do you want eggs for breakfast?

LIST 1C

1. Everybody should brush teeth before meals.

2. Once a year everything's all right.

3. Don't use up all the letter paper.

4. Anything like that's all right with me.

5. Those people outside ought to see a doctor.

6. The windows are so dirty this month I can't see.

7. Please pass the bread and butter first.

8. Don't forget to write and pay your bill.

9. Don't let the dog out of the house.

10. There's a good ballgame this afternoon.

LIST 1D

1. If you want to go it's all right.
2. Throw these old Time magazines out.
3. Do you want to wash up in the stream?
4. It's a real dark night so watch your driving.
5. I'll carry your package for you.
6. Don't you forget to shut off the water.
7. Mountain fishing is my idea of a good time.
8. Fathers used to spend more time with their children.
9. Be careful not to break the glasses.
10. I'm sorrier than you for the mistake.

LIST 1E

1. You can catch the bus across the street.
2. Tell her the news on the phone.
3. I'll catch up with you later.
4. I'll think it over and call her.
5. I don't want to go to the movies.
6. See a dentist if your tooth hurts.
7. Put that cookie back in the box.
8. You ought to stop fooling around so much.
9. Tonight that extra time's up.
10. How do you spell your name?

LIST 1F

1. Music always makes me cheer up.
2. My brother's in town for a short while.
3. We live a few miles off the main road.
4. This suit needs to go to the cleaners.
5. They ate enough green apples.
6. Have you been sick all this week?
7. Where have you been working lately?
8. There's not enough table room in the kitchen.
9. It's hard to see where he is.
10. Look out for new business.

LIST 1G

1. I'll see you right after lunch.

2. I'll see you later this afternoon.

3. White shoes are awful to keep clean.

4. You stand over there until I move.

5. There's a piece of cake left for dinner tonight.

6. Don't wait for me at the front corner.

7. It's no trouble at all to tell.

8. Hurry up with the morning paper.

9. It didn't say anything about a big rain.

10. That drugstore phone call's for you.

LIST 1H

1. Believe me it's too late.
2. Let's get that cup of coffee.
3. Let's get out of here before long.
4. I hate driving if it's at night.
5. There was water in the cellar yesterday.
6. She'll only be gone a few minutes.
7. How do you know we'll have it soon?
8. Children like candy after heavy meals.
9. No grass grows when we don't get rain.
10. They're not listed in the new phone book.

LIST 1I

1. Where can I find a place to park?
2. I like those big red apples.
3. You'll get fat by eating candy.
4. The color show's over in the Fall.
5. Why don't they paint their other walls?
6. How come you always get to go first?
7. What are you hiding under your coat?
8. I should always buy new cars.
9. What's wrong with sugar and cream in my coffee?
10. I'll wait just one minute.

LIST 1J

1. But we won't be ready to start.
2. I don't know what's wrong with the car.
3. It sure takes a sharp knife to cut meat.
4. I haven't read a newspaper since we got television.
5. The weeds are spoiling this yard.
6. Call me a little later for breakfast.
7. Do you have change for a five-dollar bill?
8. How are the things we bought?
9. I'd like some ice cream with my pie.
10. I don't think I'll have dessert.

APPENDIX E

RAW DATA

SUBJECTS' HEARING LEVELS AND DISCRIMINATION SCORES

GROUP I

Subject No.	Age (Years)	Sex	Educational Level (Yrs.)	Hearing Levels (dB)									Discrimination Scores in % Correct U.M. Test #1				MRT*			
				250	500	1k	2k	3k	4k	6k	Hz	S/N	Q	-1	-3	-5	Q	0	-3	-6 (dB)
1	46	M	12	5	5	0	0	20	15	15			100	88	79	53	97.6	88	66.4	64
2	43	F	16	0	0	0	0	0	10	0			99	93	83	73	90.4	80.8	66.4	61.6
3	27	M	12	10	5	5	5	5	5	10			99	85	64	50	95.2	78.4	83.2	68.8
4	21	F	15	-5	0	0	0	-5	5	-5			100	96	82	59	100	92.8	83.2	73.6
5	33	F	12	5	5	0	0	5	0	15			100	93	83	44	88	85.6	76	49.6
6	36	M	12	0	5	0	15	10	10	15			99	90	80	71	97.6	88	80.8	71.2
7	55	M	12	5	10	5	10	15	10	20			100	91	81	77	97.6	80.8	73.6	61.6
8	38	F	15	5	10	0	0	-5	0	0			99	93	84	54	95.2	80.8	78.4	52
9	44	M	16	-5	0	0	5	5	10	0			100	92	82	59	90.4	85.6	80.8	61.6
10	41	F	12	10	5	0	5	15	20	20			100	76	90	45	80.8	85.6	80.8	52
11	38	M	12	0	5	0	-5	0	10	15			100	84	76	70	95.2	78.4	78.4	73.6
12	31	F	18	-5	-5	-5	0	-5	5	20			100	95	85	77	92.8	88	88	83.2
13	52	F	12	5	5	0	5	10	15	15			100	78	64	26	95.2	80.8	68.8	56.8
14	34	M	12	5	5	5	15	20	20	15			99	94	81	45	88	85.6	71.2	61.6
15	30	F	13	-5	-5	-5	0	0	5	10			99	82	62	69	88	85.6	64	52
16	36	M	17	0	5	5	0	10	10	10			100	81	74	63	92.8	80.8	78.4	64

*MRT scores have been corrected for guessing

GROUP II

Subject No.	Age (Years)	Sex	Educational Level (Yrs.)	Hearing Levels (dB)									Discrimination Scores in % Correct							
													U.M. Test #1				MRT*			
				250	500	1k	2k	3k	4k	6k	Hz	S/N	Q	-1	-3	-5	Q	0	-3	-6 (dB)
1	52	F	12	0	0	15	15	25	35	20			100	81	88	63	92.8	66.4	68.8	54.4
2	48	M	9	20	15	15	20	70	70	75			98	50	50	18	71.2	47.2	40	30.4
3	54	M	16	10	10	10	35	55	60	80			98	68	47	16	90.4	54.4	49.6	25.6
4	52	M	12	5	10	10	15	40	50	45			93	57	32	23	92.8	83.2	61.6	30.4
5	47	M	10	5	10	10	10	25	65	65			98	82	42	39	88	64	56.8	37.6
6	47	M	12	5	5	5	30	60	50	60			91	47	41	18	85.6	56.8	32.8	23.2
7	47	M	10	5	10	15	25	40	50	65			100	75	46	41	85.6	52	47.2	23.2
8	46	M	12	15	15	10	20	30	30	35			98	89	81	55	92.8	76	66.4	54.4
9	37	M	10	15	10	10	10	25	70	45			98	92	86	44	78.4	68.8	59.2	47.2
10	47	M	10	10	15	5	10	85	95	90			100	53	30	14	88	56.8	44.8	23.2
11	37	M	20	5	10	10	10	30	60	55			97	83	51	16	90.4	76	76	56.8
12	54	M	12	5	10	10	30	60	70	60			100	60	37	32	85.6	47.2	28	11.2
13	38	M	18	10	10	15	10	55	65	55			99	53	37	18	88	66.4	42.4	13.6
14	27	M	14	5	5	10	30	55	45	25			99	70	48	29	95.2	83.2	61.6	42.4
15	54	M	12	10	15	15	10	45	65	60			100	76	60	30	88	61.6	54.4	35.2
16	38	M	12	0	5	15	25	55	60	55			96	47	37	1	78.4	44.8	30.4	35.2

*MRT scores have been corrected for guessing

GROUP III

Subject No.	Age (Years)	Sex	Educational Level (Yrs.)	Hearing Levels (dB)								Discrimination Scores in % Correct								
													U.M. Test #1				MRT*			
				250	500	1k	2k	3k	4k	6k	Hz	S/N	Q	-1	-3	-5	Q	0	-3	-6 (dB)
1	53	M	12	20	15	15	30	80	100	100+			94	29	18	8	68.8	32.8	11.2	1.6
2	55	M	10	15	20	20	45	55	55	70			96	49	30	0	76	59.2	44.8	13.6
3	44	M	12	0	5	25	55	70	75	75			73	13	1	0	61.6	23.2	1.6	0
4	50	M	8	15	15	20	30	60	80	85			98	46	45	4	68.8	56.8	18.4	8.8
5	55	M	12	5	5	10	70	70	65	50			85	25	26	0	80.8	61.6	30.4	6.4
6	56	M	12	15	15	20	50	75	90	95			87	43	10	0	68.8	42.4	11.2	0
7	55	M	12	15	15	20	40	50	60	80			94	42	38	2	83.2	56.8	23.2	20.8
8	46	F	17	10	10	20	45	65	70	80			94	42	21	0	76	40	25.6	0
9	53	M	12	5	10	20	30	35	45	60			100	82	81	41	85	83.2	47.2	61.6
10	52	M	12	25	20	15	30	50	50	40			100	75	56	31	90.4	73.6	64	56.8
11	56	M	13	15	20	25	25	55	60	75			95	62	24	7	88	37.6	23.2	11.2
12	52	M	12	5	20	25	35	50	40	55			100	73	50	36	80.8	61.6	54.4	35.2
13	54	M	12	15	10	20	30	25	35	50			99	75	54	37	90.4	54.4	49.6	25.6
14	50	M	9	10	15	25	35	50	60	45			96	72	46	29	73.6	49.6	35.2	20.8
15	37	M	12	5	5	10	65	90	85	90			97	37	27	5	85.6	44.8	16	0
16	42	M	8	5	15	20	45	45	35	35			95	86	72	44	90.4	76	52	52

*MRT scores have been corrected for guessing

REFERENCES

Acton, W. I. Speech intelligibility in a background noise and noise-induced hearing loss. Ergonomics, 13, 546-554 (1970).

AMA American Medical Association. Committee on Rating of Physical Impairment. Guides to the evaluation of permanent impairment: ear, nose and throat and related structures. J. Amer. Med. Assoc., 177, 99-108 (1961).

AMA American Medical Association. Committee on Rating of Physical Impairment. Ear, nose and throat and related structures. Guides to the Evaluation of Permanent Impairment, Chicago: AMA (1971).

ANSI American National Standard for Acoustical Terminology. ANSI S1.1 (1960)(R1976).

ANSI American National Standard Criteria for Background Noise in Audiometer Rooms. ANSI S3.1 (1960)(R1971).

ANSI American National Standard Methods for the Calculation of the Articulation Index. ANSI S3.5 (1969).

ANSI American National Standard Specifications for Audiometers. ANSI S3.6 (1969).

Aniansson, G. Binaural discrimination of "everyday" speech. Acta Otolaryngol., 75, 334-336 (1973).

Atherley, G. R. C. and Nobel, W. G. Clinical picture of occupational hearing loss obtained with the Hearing Measurement Scale. In D. W. Robinson (Ed.), Occupational Hearing Loss. New York: Academic (1971).

BBN Bolt, Beranek and Newman, Inc. Impact of Noise Control at the Workplace. Report No. 2671 submitted to U.S. Dept. Labor (1974).

Beranek, L. L. (Ed.), Noise and Vibration Control. New York: McGraw-Hill (1971).

Bergman, M. Hearing and Aging. Audiology, 10, 164-171 (1971).

Blumenfeld, V. G., Bergman, M., and Millner, E. Speech discrimination in an aging population. J. Speech Hearing Res., 12, 210-217 (1969).

BSI British Standards Institute. Method of Test for Estimating the Risk of Hearing Handicap Due to Noise Exposure. BS 5330 (1976).

Bullock, M. D. The effects of different reverberation times upon the intelligibility of PB words as perceived by subjects with normal hearing and subjects with sensorineural impairments and concomitant discrimination losses. Masters thesis, Ohio State Univ. (1967).

Burns, W. *Noise and Man*. Philadelphia: Lippincott (1973).

Carhart, R. Speech reception in relation to pattern of pure tone loss. *J. Speech Dis.*, *11*, 97-108 (1946).

Carter, M. A. (for AMA Council On Physical Therapy). Tentative standard procedure for evaluating the percentage of useful hearing loss in medicolegal cases. *J. Amer. Med. Assoc.*, *133*, 396-397 (1947).

CHABA, Committee on Hearing, Bioacoustics and Biomechanics. Speech Understanding and Aging, Report of Working Group 75. National Academy of Sciences, Washington, D. C. (1977).

Cooper, J. C. and Cutts, B. P. Speech discrimination in noise. *J. Speech Hearing Res.*, *14*, 322-337 (1971).

Davis, H. The articulation area and the Social Adequacy Index for hearing. *Laryngoscope*, *58*, 761-778 (1948).

Davis, H. Guide for the classification and evaluation of hearing handicap. *Trans. Amer. Acad. Ophth. and Otol.*, *69*, 740-751 (1965).

Davis, H. Hearing handicap, standards for hearing, and medicolegal rules. In H. Davis and S. R. Silverman (Eds), *Hearing and Deafness*. New York: Holt, Rinehart and Winston (1970).

Davis, H. A historical introduction. In D. W. Robinson (Ed.), *Occupational Hearing Loss*. New York: Academic (1971).

Davis, H. Some comments on "Impairment to hearing from exposure to noise" by K. D. Kryter. *J. Acoust. Soc. Amer.*, 53, 1237-1239 (1973).

De Forest, R. E. and Lierle, D. M. (for AMA Council on Physical Medicine and Rehabilitation). Principles for evaluating hearing loss. *J. Amer. Med. Assoc.*, *59*, 550-552 (1955).

Dickman, D. M. The effects of three noise backgrounds on the sentence intelligibility scores of hearing-impaired listeners. Doctoral dissertation, Univ. Maryland (1974).

Elkins,, E. Evaluation of Modified Rhyme Test results from impaired- and normal-hearing listeners. *J. Speech Hearing Res.*, *14*, 589-595 (1971).

Elkins, E. Causey, G. D., Beck, L. B., Brewer, D., and de Moll, J.B. Normal and impaired listener performance on the University of Maryland Revised CID Sentence Lists. (Unpub. report) Univ. of Maryland (1975).

EPA, U.S. Environmental Protection Agency, Occupational noise exposure regulation: Request for review and report. Fed. Register, 39, 43802-43809 (1974).

EPA, U.S. Environmental Protection Agency, Comments on Occupational Safety and Health Administration's Draft Environmental Impact Statement, "Proposed Regulation: Noise" (1975).

Fairbanks, G. Test of phonemic differentiation: The Rhyme Test, J. Acoust. Soc. Amer., 30, 596-600 (1958).

Feldman, R. and Reger, S. , Relations among hearing, reaction time, and age. J. Speech Hearing Res., 10, 479-495 (1967).

Fletcher, H. Speech and Hearing. New York: Van Nostrand (1929).

Fletcher, H. A method of calculating hearing loss for speech from an audiogram. J. Acoust. Soc. Amer., 22, 1-5 (1950).

Fletcher, H. The perception of speech sounds by deafened persons. J. Acoust. Soc. Amer., 24, 490-497 (1952).

Fletcher, H. and Galt, R. H. The perception of speech and its relation to telephony. J. Acoust. Soc. Amer., 22, 89-151 (1950).

Fowler, E. P. A Simple method of measuring percentage of capacity for hearing speech. Arch. Otolaryngol., 36, 874-890 (1942).

Fowler, E. P. Symposium on noise: (a) the percentage of capacity to hear speech and related disabilities. Laryngoscope, 57, 103-113 (1947).

French, N. R. and Steinberg, J. C. Factors governing the intelligibility of speech. J. Acoust. Soc. Amer., 19, 90-119 (1947).

Gardner, M. B. Effect of noise, system gain, and assigned task on talking levels in loud speaker communication. J. Acoust. Soc. Amer., 40, 955-965 (1966).

von Gierke, H. E. Testimony at the public hearings on the proposed standard for occupational exposure to noise, U. S. Department of Labor, OSHA (1975).

Ginnold, R. Workmen's compensation for hearing loss in Wisconsin. Labor Law J., 682-697 (Nov. 1974).

Giolas, T. G. Comparative intelligibility scores on sentence lists and continuous discourse. J. Aud. Res., 6, 31-38 (1966).

Giolas, T. G. and Epstein. Comparative intelligibility of word lists and continuous discourse. J. Speech Hearing Res., 6, 349-358 (1963).

Glorig, A. and Baughn, W. L. Basis for percent risk table. In Proceedings of the International Congress on Noise as a Public Health Problem, U.S. EPA Report 550/9-73-008 (1973).

Goetzinger, C., Proud, G., Dirks, D., and Embrey, J. A study of hearing in advanced age. Arch. Otolaryngol., 73, 662-673 (1961).

Guild, E. Personal protection. In Industrial Noise Manual, (2nd Ed.). Detroit: AIHA (1966).

Guilford, J. P. Fundamental Statistics in Psychology and Education, (4th Ed.). New York: McGraw-Hill (1965).

Harris, C. M. (Ed.), Handbook of Noise Control. New York: McGraw-Hill (1957).

Harris, J. D. Pure-tone acuity and the intelligibility of everyday speech. J. Acoust. Soc. Amer., 37, 824-830 (1965).

Harris, J. D., Haines, H. L., and Myers, C. K. A new formula for using the audiogram to predict speech hearing loss. Arch. Otolaryngol., 63, 158-176 (1956), also Erratum. 64, 477 (1956).

Harris, J. D., Haines, H. L., and Myers, C. K. The importance of hearing at 3kc for understanding speeded speech. Laryngoscope, 70, 131-146 (1960).

Harris, J. D., Haines, H. L., Kelsey, P. A., and Clack, T. D. The relation between speech intelligibility and the electroacoustic characteristics of low fidelity circuitry. J. Aud. Res., 1, 357-381 (1961).

High, W. S., Fairbanks, G., and Glorig, A. Scale for self-assessment of hearing handicap. J. Speech Hearing Dis., 29, 215-230 (1964).

Hirsh, I. J., Reynolds, E. G., and Joseph, M. Intelligibility of different speech materials. J. Acoust. Soc. Amer., 26, 530-538 (1954).

House, A. S., Williams, C. E., Hecker, M. H. L., and Kryter, K. D. Articulation-testing methods: Consonantal differentiation with a closed-response set. J. Acoust. Soc. Amer., 37, 158-166 (1965).

ISO International Organization for Standardization, ISO Standard 1999 - Acoustics - Assessment of Occupational Noise Exposure for Hearing Conservation Purposes (1971).

ISO International Organization for Standardization, Recommended Methods for Measuring the Intelligibility of Speech. First draft proposal of ISO/TC 43/SC 1/WG 12 (Nov. 1975).

Jerger, J. Audiological findings in aging. Advances in Oto-rhinolaryngology, 20, 115-124, Basel (1973).

Klockhoff, I., Drettner, B., Hagelin, K. W., and Lindholm, L. A method for computerized classification of pure tone screening audiometry results in noise-exposed groups. Acta Otolaryngol., 75, 339-340 (1973).

Knudsen, V. O. The sensibility of the ear to small differences of intensity and frequency. Phys. Rev., 21, 84-102 (1923).

Korn, T. S. Effect of psychological feedback on conversational noise reduction in rooms. J. Acoust. Soc. Amer., 26, 793-794 (1954).

Kreul, E. J., Nixon, J. C., Kryter, K. D., Bell, D. W., Lang, J. S., and Schubert, E. D. A proposed clinical test of speech discrimination. J. Speech Hearing Res., 11, 536-552 (1968).

Kryter, K. D. Effects of ear protective devices on the intelligibility of speech in noise. J. Acoust. Soc. Amer., 18, 413-417 (1946).

Kryter, K. D. Hearing Impairment for speech. Arch. Otolaryngol., 77, 598-602 (1963).

Kryter, K. D. The Effects of Noise on Man. New York: Academic (1970).

Kryter, K. D. Speech communication. In H. P. Van Cott, and R. G. Kinkade (Eds.), Human Engineering Guide to Equipment Design. Washington, DC: Amer. Inst. for Res. (1972).

Kryter, K. D. Impairment to hearing from exposure to noise. J. Acoust. Soc. Amer., 53, 1211-1234 (1973).

Kryter, K. D. Hearing impairment and damage risk from noise exposure. Testimony at the public hearings on the proposed standard for occupational exposure to noise. U.S. Department of Labor, OSHA (1975).

Kryter, K. D. and Whitman, E. C. Some comparisons between rhyme and PB-word intelligibility tests. J. Acoust. Soc. Amer., 37, 1146 (1965).

Kryter, K. D., Williams, C., and Green, D. M. Auditory acuity and the perception of speech. J. Acoust. Soc. Amer., 34, 1217-1223 (1962).

Kuzniarz, J. J. Hearing loss and speech intelligibility in noise. In Proceedings of the International Congress on Noise as a Public Health Problem. U.S. EPA Report 550/9-73-008 (1973).

Liden, G. Medicolegal problems in the assessment of deafness. Acta Otolaryngol., Suppl. 158 (1960).

Lierle, D. M. (for the Subcommittee on Noise of the AAOO Committee on Conservation of Hearing). Guide for the evaluation of hearing impairment. Trans. Amer. Acad. Ophthal. Otolaryng., 63, 236-238 (1959).

Lindeman, H. E. Relation between audiological findings and complaints by persons suffering noise-induced hearing loss. J. Amer. Indus. Hyg. Ass., 32, 447-452 (1971).

Littler, T. S. Physics of the Ear. London: Pergamon (1965).

MacKeith, N. W. and Coles, R. R. A. Binaural advantage in hearing of speech. J. Laryng. Otol., 85, 213-232 (1971).

Mayer, C. The perception of speech in noise as a function of age. Doctoral dissertation, City University, N. Y. (1975).

Merluzzi, F. and Hinchcliffe, R. Threshold of subjective auditory handicap. Audiol., 12, 65-69 (1973).

Meyer, A. F. Testimony at the public hearings on the proposed standard for occupational exposure to noise. U.S. Department of Labor, OSHA (1975).

Miller, G. A. The masking of speech. Psychol. Bull., 44, 105-129 (1947).

Millin, J. P. The effect of small room reverberation on discrimination tests. Doctoral dissertation, Case Western Reserve Univ. (1968).

Miner, R. and Danhauer, J. L. Articulation functions for normal and hearing impaired subjects on the MRT and SSIT in multitalker noise. Paper presented at the 91st meeting of the Acoustical Society of America. Washington, DC (1976).

Moncur, J. P. and Dirks, D. D. Binaural and monaural speech intelligibility in reverberation. J. Speech Hearing Res., 10, 186-195 (1967).

Mullins, C. J. and Bangs, J. L. Relationships between speech discrimination and other audiometric data. Acta Otolaryngol., 47, 149-157 (1957).

Murry, T. and Lacroix, P. G. Speech discrimination in noise and hearing loss at 3000 Hz. Nav. Sub. Med. Res. Lab. Report No. 719, Groton, Conn. (1972).

Myers, C. K. and Angermeier, C. Hearing loss at 3kHz and the Chaba "Proposed Clinical Test of Speech Discrimination in Noise." Nav. Sub. Med. Res. Lab. Report No. 720 Groton, Conn. (1972).

Nabelek, A. K. and Pickett, J. M. Reception of consonants in the classroom as affected by monaural and binaural listening, noise, reverberation, and hearing aids. J. Acoust. Soc. Amer., 56, 628-639 (1974a).

Nabelek, A. K. and Pickett, J. M. Monaural and binaural speech perception through hearing aids under noise and reverberation with normal and hearing-impaired listeners. J. Speech Hearing Res., 17, 724-739 (1974b).

Nett, E. M., Doerfler, L. G., and Matthews, J. The relationships between audiological measures and handicap. Unpub. report, Univ. of Pittsburg Sch. Med. (1959).

Newby, H. A. Audiology, (2nd ed.). New York: Appleton-Century-Crofts (1964).

Niemeyer, W. Speech discrimination in noise-induced deafness. Int. Aud., 6, 42-47 (1967).

NIOSH, U.S. Department of Health, Education, and Welfare, National Institute for Occupational Safety and Health, Criteria for a recommended standard: Occupational exposure to noise. HSM 73-1001 (1972).

Northern, J. L. and Hattler, K. W. Evaluation of four speech discrimination test procedures on hearing impaired patients. J. Aud. Res., suppl. 1 (1974).

OSHA, U.S. Department of Labor, Occupational Safety and Health Administration, Draft Environmental Impact Statement: Proposed Regulation, Noise. (1975).

Pearsons, K. S., Bennett, R. L., and Fidell, S. Speech levels in various environments. Bolt, Beranek and Newman, Inc., BBN Report No. 3281 (1976).

Pickett, J. M. Effects of vocal force on the intelligibility of speech sounds. J. Acoust. Soc. Amer., 28, 902-905 (1956).

Pickett, J. M. Limits of direct speech communication in noise. J. Acoust. Soc. Amer., 30, 278-281 (1958).

Quiggle, R. R., Glorig, A., Delk, J. H. and Summerfield, A. B. Predicting hearing loss for speech from pure-tone audiograms. Laryngoscope, 67, 1-15 (1957).

Quist-Hanssen, S. and Steen, E. Observed and calculated hearing loss for speech in noise-induced deafness. Acta Otolaryng., suppl. 158, 277-281 (1960).

Risberg, A. and Martony, J. A method for classification of audiograms. In G. Fant (Ed.), Proceedings of the Symposium on Speech Communication Ability and Profound Deafness, Washington, DC: A. G. Bell Assoc. for the Deaf (1972).

Robinson, D. W. (Ed.) Occupational Hearing Loss. New York: Academic (1971).

Ross, M. Huntington, D. A., Newby, H. A., and Dixon, R. F. Speech discrimination of the hearing-impaired individual in noise. J. Aud. Res., 6, 47-72 (1965).

Rudmose, W. Hearing loss resulting from noise exposure. In C. M. Harris (Ed.), Handbook of Noise Control. New York: McGraw-Hill (1957).

Sabine, H. J. Less Noise Better Hearing. Chicago: Celotex Corp. (1950).

Sabine, P. E. On estimating the percentage loss of useful hearing. Trans. Amer. Acad. Ophthal. Otolaryng., 46, 179-196 (1942).

Scott, W. A. and Wertheimer, M. Introduction to Psychological Research. New York: John Wiley & Sons (1964).

Shower, E. G. and Biddulph, R. Differential pitch sensitivity of the ear. J. Acoust. Soc. Amer., 3, 275-287 (1931).

Silverman, S.R. and Hirsh, I.J. Problems related to the use of speech in clinical audiometry. Ann. Otol. Rhinol. Laryng., 64, 1234-1244 (1955).

Silverman, B., Bilger, R. C., Hanley, T. D., and Steer, M. D. The relative intelligibility of male and female talkers. J. Educ. Psychol., 44, 418-428 (1953).

Simonton, K. M. and Hedgecock, L. D. A laboratory assessment of hearing acuity for voice signals against a background of noise. Ann. Otol., 62, 735-747 (1953).

Steinberg, J. C. and Gardner, W. B. On the auditory significance of the term hearing loss. J. Acoust. Soc. Amer., 11, 270-277 (1940).

Suter, A. H. and von Gierke, H. E. Evaluation and compensation of occupational hearing loss in the United States. In Proceedings of the Congress, Man and Noise. Turin, Italy: Edizioni Minerva Medica (1975).

Thomas, W. G., Testimony at the public hearings on the proposed standard for occupational exposure to noise. U.S. Department of Labor, OSHA (1975).

Tillman, T. W. and Carhart, R. An expanded test for speech discrimination utilizing CNC monosyllabic words. USAF School of Aerospace Med., Brooks AFB, SAM-TR-66-55 (1966).

UK-Member Body of ISO/TC 43/SC1, Proposal for a revision of ISO 1999 - Acoustics - Assessment of Occupational Noise Exposure for Hearing Conservation Purposes (1975).

U.S. Air Force, Aerospace Medicine, Hazardous Noise Exposure. AFR 161-35, Attachment 9 (27 July 1973).

Utley, J. The relationship between speech sound discrimination and percentage hearing loss. J. Speech Hearing Dis., 9, 103-113 (1944).

VA, U.S. Veterans Administration, Rating procedure relative to specific issues: 50.07 rating of hearing impairment. Dept. Med. and Surgery Manual, M21-1 (1976).

Ward, W. D., Fleer, R. E., and Glorig, A. Characteristics of hearing losses produced by gunfire and by steady noise. J. Aud. Res., 21, 325-356 (1962).

Webster, J. C. Important frequencies in noise-masked speech. Arch. Otolaryngol., 80, 494-504 (1964).

Webster, J. C. Effects of noise on speech intelligibility in W. D. Ward and J. E. Fricke (Eds.), Noise as a Public Health Hazard, ASHA Report No. 4. Washington, DC: American Speech and Hearing Association (1969).

Webster, J. C. Compendium of speech testing materials and typical noise spectra for use in evaluating communications equipment. Naval Electronics Laboratory Center, Human Factors Technology Division, Tech. Doc. 191 (13 Sept 1972).

Webster, J. C. The effects of noise on the hearing of speech. In Proceedings of the International Congress on Noise as a Public Health Problem, U.S. EPA Report 550/9-73-008 (1973).

Webster, J. C. and Klumpp, R. G. Effects of ambient noise and nearby talkers on a face-to-face communication task. J. Acoust. Soc. Amer., 34, 926-941 (1962).

Webster, J. C., Davis, H., and Ward, W. D. Everyday speech intelligibility. J. Acoust. Soc. Amer., 38, 668 (1965).

Winer, B. J. Statistical Principles in Experimental Design (2nd ed.). New York: McGraw-Hill (1971).